Research article

An analysis of Koreans’ attitudes towards migrants by application of algorithmic approaches

Seungwoo Han

Division of Global Affairs, Rutgers University, United States

ARTICLE INFO

Keywords:
Multiculturalism
Migrants
Machine learning
LightGBM
SHAP
South Korea

ABSTRACT

Exclusivist behaviours toward migrants manifest not only from poor receptivity to multiculturalism, but also from competition for resources and opportunities such as job opportunities. This study analyses the socio-economic background of Koreans by categorizing them in terms of their receptivity to multiculturalism and their stances on the job debate. For its analytical methodology, this study proposes an analytical approach that combines survey data from questionnaires with machine-learning algorithms. This study applies light gradient boosting machine (LightGBM) and SHapley Additive exPlanations (SHAP) to estimate the importance of the studied variables and interpret the estimate results. According to the results of this study, contact, rapport, experience, and familiarity with other cultures promoted receptivity, but competition over job opportunities led to realistic threats.

1. Introduction

Since the mid-1990s, Koreans have slowly acclimated to co-existing with migrants. South Korea (henceforth Korea) has adopted policies to promote receptivity toward other cultures, and mitigate exclusivist attitudes against migrants, through political and educational interventions that seek to overcome Korea’s poor multiculturalist foundations (Lee, 2012; Kim, 2012a, 2015; Park et al., 2013; Park, 2014). However, exclusivist behaviours toward migrants manifest not only from poor receptivity to multiculturalism, but also from competition for resources and opportunities such as job opportunities. This study explores the different viewpoints of Koreans on these two aspects through comparative analysis.

This study analyses the socio-economic background of Koreans by categorizing them in terms of their receptivity to multiculturalism and their stances on the job debate. For its analytical methodology, this study proposes an analytical approach that combines survey data from questionnaires with machine learning algorithms. First, this study overcomes the limitations of conventional analytical approaches by applying the Light Gradient Boosting Machine (LightGBM) and SHapley Additive explanations (SHAP) algorithms. In addition, this study uses data from the 2018 National Multicultural Acceptability Survey conducted by the Ministry of Gender Equality and Family, and data from various sources to discuss and analyse the results.

This article comprises five sections. Section 2 discusses Korea’s multiculturalism and theoretical background. Section 3 discusses the study’s methodology. Section 4 covers data analysis and discussion. Section 5 concludes the study and presents some implications.

2. Multiculturalism in South Korea and theoretical background

Korea has a short history of accepting migrants from other countries compared to countries with long immigration histories, because there was less active interaction with other countries or ethnic groups for geopolitical reasons until 1990s. Therefore, Korean society has not handled key issues of migration, such as public order, migration regulation, and illegal migration, deliberately. In addition, Korea does not have historical problems stemming from having overseas colonies or a massive influx of citizens of (former) colonies; because of its short migration history, religious or racial conflicts and right-wing extremism have not fully surfaced. However, considering the fact that Korea’s multi-racial and multi-cultural manifestation is following other countries’ experiences in a compressed timeframe, it is very likely that racial and religious conflicts will soon become serious issues in Korea.

The history of migration to Korea began with the Korean-Chinese, after the establishment of diplomatic relations between China and Korea in 1992 (Lee, 2003; Kim, 2009a, 2009b; Lee, 2008; Freeman, 2011; Lee, 2012). Korean society has manifested an ambivalent response to-
wards migrants: a mixed sentiment of tolerance and wariness in reality, with a mixed view of multiculturalism and assimilation in its social discourse (Be'langer et al., 2010; Kim, 2011; Kim, 2012a; Lee, 2012; Kim, 2013; Park, 2014; Kim, 2015). Countries with comparatively long histories of co-existence with migrants have deliberated methods of co-existing with migrants and implemented various policies (Hammer, 1985; Cuccioletta, 2001; Hing, 2004; Knowles, 2016). Korean society, on the other hand, is very far from the stage where it discusses how it can embrace a new influx of migrants after co-existing with ethnic minorities or migrants for a long time (Lee, 2009). Recently, a large number of migrants have moved to Korea while the country has been struggling with stagnant economic growth and its weakened ability to integrate citizens due to increasing polarization, inequality, and exclusion (Han, 2022a, 2022b). In particular, there is a dearth of experience of coexisting with various ethnic groups since the early-modern period. Thus, Korean society’s limitations in multicultural discourse originate from the lack of experience of coexistence with minor ethnic groups. In this light, Korea’s perception of migrants derives not from a long-history of co-existence and experiences, but from the informative and educational policies of its central government and press (Lee, 2012; Kim, 2012a, 2015; Park et al., 2013; Park, 2014).

Policy and education driven by the government are the primary drivers of the effort to promote multiculturalist receptivity and mitigate exclusionary attitudes toward migrants. In Korean society, a mono-ethnic ideology has been dominant (Lee, 2009; Kim, 2013, 2015). Educational programs for overcoming exclusionary attitudes based on the mono-ethnic ideology have targeted the public, teachers, and students under the ultimate goal of respecting the differences between Koreans and migrants and building mutual understandings.

Because Korea has had a very short history of co-existence with migrants, prior studies focus on Koreans’ perceptions and attitudes towards other cultures (Kwon, 2009; Jo, 2009; Song, 2010; Lee, 2010; Oh and Roh, 2014; Kwak et al., 2019). Some prior studies focus on analysing the socio-economic factors that caused differences in receptivity towards migrants (Won, 2011; Oh et al., 2014; Ha et al., 2014; Kwak et al., 2019). This study takes the perspective that the discussion on exclusionary attitudes toward migrants should consider economic opportunities as well.

Prior studies that attempt to identify the backdrop receptivity towards migrants by occupational status prove that receptivity towards them varied by employment types (McLaren, 2003; Scheve and Slaughter, 2006; Kaya, 2012; Callens et al., 2015; Harris et al., 2018; Gorodzeisky and Semyonov, 2019; Buehler et al., 2020). However, these previously-mentioned studies do not clarify the relationship between socio-economic backgrounds and their positive or negative attitudes toward migrant employment opportunities. While multicultural receptivity correlates with a person’s stance on migrant employment opportunities, they are separate issues. By separating the two issues, present study seeks to present a new perspective on exclusionary attitudes toward migrants.

Countries with strong mono-ethnic ideologies do not always manifest an exclusionary attitude toward migrants. Exclusionary attitudes toward a certain group exist in any countries where there are large economic inequalities or high unemployment (Powell, 2017; Rzepnikowska, 2019).
In countries with long histories of co-existence with ethnic minorities and migrants, the manifestation of exclusionary attitudes toward migrants is not necessarily due to low receptivity towards multiculturalism (McLaren, 2001; Calavita, 2005; McGhee, 2005; Fussell, 2014). An exclusionary attitude toward migrants does not stem purely from ethnic or cultural attributes; it can also rise in the minds of citizens who are in direct or indirect competition with migrants at a time when there is a shortage of jobs, or a growing economic gap (Park, 1998, 2000; Powell, 2017; Rzepnikowska, 2019; Gover et al., 2020). Therefore, individuals’ stances toward migrants vary based on their socio-economic status and occupation (Callens et al., 2015; Harris et al., 2018; Gorodzeisky and Semyonov, 2019; Buehler et al., 2020). Without a doubt, exclusionary attitudes towards migrants intensify when the discourse added nationalism (Schmidt and Quandt, 2018).

Intergroup relations refers to the general way that people in a certain social group or category think of, feel about, behave toward, or interact with individuals from a different social group (Hogg, 2013). Similarly, inter-ethnic relations in a multicultural society refer to how one ethnic group views and interacts with other ethnic groups. According to Realistic Group Conflict Theory, which highlights potential economic conflicts between groups, competition over scarce resources may lead to intergroup conflict, and exclusionary attitudes toward migrants originate from people’s disapproval of the economic threat posed by migrant labour (Campbell, 1965; Esses et al., 1998; Baumeister and Vohs, 2007). According to Social Identity Theory, people desire a positive social identity that leads them to establish positive traits unique to their own group, in comparison to other groups (Tajfel and Turner, 1986). This manifested in hostility, intolerance, and prejudice toward other groups in their daily lives. According to the theory, the more typical the intergroup behaviour is, the more consensual an attitude it displays toward out-groups within the in-group, and the more homogeneous they consider the members of the out-group to be. The intergroup comparison focuses on establishing the positive attributes of the in-group.

Multicultural receptivity assigns certain attributes to a person’s group and other groups based on the person’s educational and cultural experience with migrants. A different logic governs the person’s stance on jobs, however, since it is associated with the person’s economic activity and livelihood instead. Nonetheless, it is difficult to distinguish the two clearly, and in reality, conflicts and prejudices between groups are complicated and assume complicated patterns (Zarate et al., 2004; Hogg, 2013). For example, even if an individual has a high tolerance for migrants based on their education and experiences of diverse cultures, that person may adopt a different stance if they have to compete over jobs or develop a competitive mind-set.

Given this perspective, Integrated Threat Theory provides a variety of approaches to the causes and interpretations of intergroup conflicts (Stephan and Stephan, 1996; Stephan et al., 1999a, 1999b; Cürçiu et al., 2007). The theory postulates the following four perspectives: realistic threats, symbolic threats, intergroup anxiety, and negative stereotypes. First, realistic threats derive from the practice of competition. This type of threat manifests, not from stereotypes or misbeliefs, but because of competitive relationships over employment and resources in the group. Therefore, when competitions arise over scarce resources between certain in-groups and out-groups, the group members experience a recognizable, realistic conflict based on competition (Stephan and Stephan, 1996; Esses et al., 1998; Hogg, 2013). Second, the theory presents symbolic threats, which stem from differences in moral values, norms, social standards, beliefs, and attitudes. In other words, symbolic threats
arise when the normative values and order of an in-group are broken or come into conflict due to new values and symbols from out-groups (Stephan and Stephan, 1996; Stephan et al., 1999a). Third, intergroup anxiety refers to the negative emotions and conflicts that arise when members of a group expect, or actually experience, interactions with members of other groups. That is, anxiety is based on the negative emotions that exist in the relationships between in-group and out-group members (Stangor et al., 1991; Stephan et al., 1999a, b). The last perspective presented by the theory is negative stereotypes. It refers to the negative beliefs, expectations, or attitudes of one group toward other groups. Unlike the other perspectives, it is an imaginary or anticipated intergroup anxiety, because it is based on negative beliefs (Taifel, 1981; Hogg, 2013; Stephan et al., 1999a, 1999b).

Integrated Threat Theory postulates that a particular threat inherent in the multiple groups (as opposed to a single group) that a person belongs to is the main factor in forming a negative perception and prejudice toward an out-group. The theoretical supposition of Integrated Threat Theory is very useful, not only for explaining the entire in-group’s discriminative and exclusionary attitude towards out-groups, but also very effective in accounting for individual and personal variations in the in-group’s perceptions towards out-groups. The theory is helpful in this study’s comparative analysis of individuals’ socio-economic backgrounds and their relationship to multicultural receptivity and the job debate.

3. Methodology

Traditional approaches in social sciences begin with the formulation of hypothesis, creation of formal models that represent the underlying causalities, and then by the test of these models on the available data. On the other hand, machine learning algorithms try to find hidden patterns from data and explain cases. To reiterate, for the traditional approaches, researchers have some expectations about causalities and try to verify these expectations empirically. On the other hand, machine learning algorithms describe certain case (certain social phenomenon) based on data without formulating causality.

Although advantages of algorithmic approaches have been appreciated (Mullainathan and Spiess, 2017), compared with the overwhelming majority of traditional approaches focusing on causality, the volumes of descriptive studies are limited in social science studies (Gerring 2012; Grimmer, 2015). Although they have been increasingly applied, still algorithmic approaches such as machine learning algorithms are rather considered as exotic approaches to social science. The dismissal of approaches to describing social phenomena based on data-driven methods is ironic. This is because much of empirical works of social sciences and theories that social scientists construct are direct products of description. Therefore, although descriptive inference based on data-driven methods is still not common in social science research, they can be useful approaches for understanding social phenomena based on data.

Figure 3. Unemployment rates and migrant populations. Sources: (a) Statistics Korea (https://kosis.kr/statHtml/statHtml.do?orgId=101&amp;tblId=DT_1DA7102S&amp;conn_path=13). (b) Ministry of Employment and Labour (https://eboard.moel.go.kr/indicator/detail?menu_idx=75), Ministry of Justice (https://www.mois.go.kr/fri/bbs/type001/commonSelectBoardArticle.do?bbsId=BBSMSTR_000000000014&amp;nttId=80781).

Figure 4. Migrant populations and overseas travel experiences by region. Sources: (a) Ministry of Justice (https://www.mois.go.kr/fri/bbs/type001/commonSelectBoardArticle.do?bbsId=BBSMSTR_000000000014&amp;nttId=80781). (b) Statistics Korea (https://kosis.kr/statHtml/statHtml.do?orgId=101&amp;tblId=DT_1YL20251&amp;conn_path=12).
Algorithmic approaches attract attention due to the usage of a large volume of a variety of data. Here, we need to pay attention to the variety of data, not the size of data. That is, the data-based approach is useful not only for unstructured data such as text or images but also for structured data such as surveys. Recently, we can find examples in various fields in social science. For instance, by applying decision tree and random forests algorithms, Hegelich (2016) analyses extreme shifts in the US budget and their connection to the attention of political actors based on conventional data. Rudra (2014) applies a clustering algorithm using macro-level data to describe the welfare state typology of developing countries. Han (2022b) describes the spatial form of inequality by collecting structured data provided by the government and applying a clustering algorithm. Oh et al. (2019) analyse variables explaining mental depression by applying various algorithms to national health surveys. Moreover, several studies apply decision tree and tree-based models to surveys to estimate inequality of societies (Brunori et al., 2019; Brunori and Neidhofer, 2020; Han, 2022c; Lefranc and Kundu, 2020).

These aforementioned studies attempt to describe certain social phenomena by applying machine learning algorithms to conventional data rather than the traditional causality inference-based approach or analysis using the vast volume of unstructured data. In this respect, the current study aims to describe the attitudes of Koreans towards migrants by applying LightGBM, which maintains the unique characteristics of the decision tree algorithm and overcomes its drawbacks pointed out (Ke et al., 2017), and the SHAP algorithm, which is suggested as a useful tool to explain the results though tree-based algorithms (Lundberg and Lee, 2017a). Although this proposal is made in the data science discipline, not in social science, and it is mainly applied to data analysis projects rather than social science researches, it is meaningful in an attempt to analyse conventional survey data and interpret the results as a novel approach. We now turn to the research methods applied in this study.

3.1. LightGBM

In order to estimate the effects of people’s socio-economic backgrounds on their receptivity towards multiculturalism via questionnaire surveys, prior studies have conducted regression analysis as follows (Kaya, 2012; Oh et al., 2014; Callens et al., 2015; Harris et al., 2018; Gorodzeisky and Semyonov, 2019; Kwak et al., 2019; Buehler et al., 2020).

$$y_i = \beta X_i + \varepsilon$$ (1)

The variables (X) representing a person’s demographic and socio-economic background mostly consist of categorical variables, except for the continuous variables such as age. Regression analysis converts categorical variables into dummy variables. Surveys measure attitudes, beliefs, opinions, and specific information about individual respondents. That is, the values subsume the effects and the direction of the categorical variable. However, converting a categorical variable into dummy variables erases certain information about the values of categorical variable. In contrast, the decision tree algorithm that underlies LightGBM’s basic model does not need to convert the categorical variable into dummy variables (Murphy, 1998; Murphy, 2012).

The decision tree is classified as a nonparametric machine learning method because it is not based on a probability density unlike parametric machine learning algorithms (Murphy, 1998). Moreover, it does not make statistical assumptions about parameters and does not imply that the underlying relationship between variables is linear (Hastie et al., 2009; Murphy, 2012). These advantages enable the use...
of limited observations even when huge amounts of data are not available.

The decision tree algorithm learns the rules inherent in the data, and infers prediction rules for the entire domain by identifying if/else statements for each variable category and repeatedly dividing them (Hastie et al., 2009; Meng et al., 2016). It can handle continuous and categorical variables simultaneously for regression or classification. However, the greatest shortcoming of the decision tree algorithm is overfitting. Overfitting results from a greedy nature of the tree splitting. Due to the hierarchical nature of the tree growth process, small changes to the data entered can completely transform the overall structure of the tree, and an initial error can greatly affect the rest of the tree (Murphy 2012). For this reason, the decision tree algorithm is by nature unstable.

Ensemble learning is a process designed to overcome shortcomings such as overfitting (Hastie et al., 2008; Ke et al., 2017). Ensemble learning is a method that can elicit more accurate and stable predictions by using several weak learners and combining their predictions. This study employs LightGBM as a model that makes up for the shortcomings of decision tree algorithms. LightGBM is an algorithm that uses a boosting framework. Boosting is a learning method that reduces errors by assigning added weights to accurately predict data that were inaccurately predicted during the ensemble learning’s decision tree process that combined several weak learners (Hastie et al., 2008; Maloney et al., 2012; Chen and Guestrin, 2016; Bonaccorso, 2018). Eq. (2) explains its core as a forward stage-wise additive model, which is a basic boosting algorithm approach.

\[
G(x) = \sum_{m=1}^{M} \alpha_m G_m(x) \tag{2}
\]

This model creates a gradient weak learner, \(G_m(x)\), \(m = (1, 2, ..., M)\), with \(G_m(x)\) receiving the final value. \(\alpha_m(x)\), \(m = (1, 2, ..., M)\) is the added weight for each weak learner, which continues to be updated to boost the performance for the next step. That is, it sequentially generates continuous learners from \(G_1(x)\) to \(G_M(x)\), and obtains the final output through an update that minimizes the error function. Among boosting models, the models that minimize the error function through gradient descent are called Gradient Boosting Decision Trees (GBDT), and LightGBM belongs in the GBDT category (Ke et al., 2017).

Describing the method of LightGBM in detail, LightGBM’s splitting method is called Gradient-based One-Side Sampling (GOSS). GOSS’s basic approach is analogous to Eq. (2). LightGBM uses the gradient from each data instance as a proxy for sample weight (Ke et al., 2017). That is, when a sample connects to a low gradient, it is possible that the training error for that instance is small and already well-trained. Thus, after the calculation is complete, there is an option to drop the samples with the smallest gradients in order to save calculation time when looking for the best splits. However, actually dropping every sample with a gradient below a certain threshold creates the possibility of having already changed the sample distribution, which decreases the explanatory power of the model. Therefore, GOSS maintains samples with steep gradients while random-sampling any instances with slight gradients. To make up for the effect from partially selecting samples with small gradients, GOSS introduces a constant multiplier for the data instances with small gradients while calculating the information gain. GOSS first sorts data instances based on the absolute value of the sample’s gradient, then selects the top a*100% of instances. Next, it does a random-sampling of b*100% instances with a slight gradient. Then, GOSS amplifies the instances with small gradients by the constant \(\frac{1}{b-a}\). In this manner, GOSS shifts the focus of the model onto the instances that are not fully trained.

In GOSS, the LightGBM model uses the variance gain \(\tilde{V}_j(d)\) for variable \(j\) and the split point, as given below (Ke et al., 2017).

\[
\tilde{V}_j(d) = \frac{1}{n} \left( \frac{\sum_{x \in L_j} g_i + \frac{1-a}{b-a} \sum_{x \in R_j} g_i}{n_l(d)} \right)^2 + \frac{\sum_{x \in R_j} g_i + \frac{1-a}{b-a} \sum_{x \in L_j} g_i}{n_r(d)} \right)^2 \tag{3}
\]

Here, \(A\) refers to data instances with steep gradients, and \(B\) refers to instances with slight gradients. \(g_i\) is the gradient of each sample, and \(A_1, B_1, A_2, B_2\) are each defined as follows.

![Figure 7. Industry structure by region (2018). Source: Statistics Korea (https://kosis.kr/publication/publicationThema.do?pubcode=GS).](image)

![Figure 8. Farming population and aging. Source: Ministry of Agriculture, Food, and Rural Affairs (https://www.index.go.kr/potal/main/EachDtlPageDetail.do?id=x_cd=2745).](image)
SHAP is an approach based on game theory, and borrowed from a concept designed by Shapley (1953). Proposed by Lundberg and Lee (2017a), SHAP values are numbers that indicate how much each variable contributed to the estimated outcome, with the goal of interpreting an algorithm’s estimated outcome. SHAP elaborates on the approach of additive feature attribution methods. The goal of SHAP is to calculate each variable’s attribution to the estimation, and to account for the result in a consistent manner. The explanatory model can explain such an approach as the linear function of a binary variable, as follows.

\[
g(z) = \phi_0 + \sum_{i=1}^{M} \phi_i z_i
\]

In this equation, \(g(z)\) is a local surrogate model of the original model \(f(x)\). Using this allows us to interpret the original model. In the case of \(z = \{0,1\}^M\), \(M\) refers to the number of explanatory variables, \(z_i\) refers to the observation status of the variables, and \(\phi_i\) refers to the attribution to the expected output for a given \(S\). \(\phi_i\) is defined as the added weight that calculates the number of subset \(S\)’s permutations, and \(f_s(S)\) is the expected output for a given \(S\). Eq. (7) allows for the calculation of the global importance of each variable. The size of this outcome determines the importance of variable \(i\).

\[
l_i = \frac{1}{N} \sum_{i=1}^{N} |\phi_i|
\]

We can estimate the importance of each variable via Eq. (7), but this does not illustrate how it works for the output. In order to discover that, we use the SHAP summary plot. The SHAP summary plot utilizes \(\phi_i\) to deliver every aspect of the variable’s importance in an intuitive visual manner (Lundberg and Lee, 2017b).

### 3.3. Data

This study utilizes data from the 2018 public survey on multicultural receptivity provided by the Ministry of Gender Equality and Family, Republic of Korea. They conducted the survey nationwide, with 4000 respondents which included males and females between the ages of 19 and 74. The sample was extracted using the multi-stage area probability sampling method, and the survey was conducted using a person-to-person interview method.

The survey included 21 variables that reflected the person’s socio-economic background. The dependent variables were 1) a numerical value for multicultural receptivity and 2) pros and cons about giving preferential employment opportunities to Korean citizens during the job shortage. The survey that quantified receptivity towards multiculturalism was comprised of the following sub-components of multicultural receptivity: cultural openness, national identity, stereotypes, discrimination, unilateral expectation of assimilation, rejection/avoidance sentiment, dual evaluations, and willingness to exchange behaviours and act as global citizens (see Supplementary Material for a description of the data, survey and variables).

### 4. Results of algorithmic analysis and discussion

This section is about results of algorithmic analysis and discussion using data from various sources. The first data analysis examines how...
Koreans have different positions on the two dependent variables and explain the reasons by examining various data.

### 4.1. Results of algorithmic analysis

There are three things to note when interpreting SHAP plots. 1) The vertical order indicates the order of importance for each variable. 2) The coloured indicates high values of the corresponding variables, and the colour blue indicates low values. 3) The horizontal axis indicates the effect of the corresponding value on the output. In Figure 1, the horizontal axis refers to the person’s receptivity towards multiculturalism; the area to the left of 0 indicates a negative attitude towards multiculturalism, while the area to the right of 0 indicates a positive attitude. Figure 1 illustrates the results from analysing the multicultural receptivity data. The performance of the model, as measured by RMSE, is 0.15889.

Figure 1 reveals that the variables estimated to be the most important for multicultural receptivity are region, age, relation, education, and urban living. In terms of region, living in the capital and the Chungcheong/Gangwon region promoted receptivity towards multiculturalism, whereas living in the Honam-Janu region inhibited receptivity towards multiculturalism. In terms of age, being younger increased receptivity towards multiculturalism, which gradually decreased with age. In terms of relationships, having family, relatives, friends, and colleagues promoted receptivity, whereas ‘None’ reduced receptivity. In terms of educational levels, higher levels of education promoted receptivity, whereas lower level of education inhibited it.

In terms of urban living, living in a rural area or small/medium-sized city promoted receptivity, whereas living in a metropolitan city reduced it. In terms of overseas traveling, more experience with overseas travel promoted receptivity. In terms of occupational position, higher values (such as temporary daily labour or unpaid family business) promoted receptivity, whereas low values (such as full-time workers or business owners with employees) inhibited it. In terms of gender, being male promoted receptivity, while being female inhibited it. In terms of income, the area to the right of 0 indicates approval. The model’s performance in-

### 4.2. Discussion

In the algorithmic analysis, some variables had very different results in Figure 2 than they had in Figure 1. We can compare the two Figures and focus on the key variables that exhibited the most contrasting differences: age, region, relationship, and occupational status. First, younger respondents were more flexible about traditional culture and conservative values, which promoted multicultural receptivity; at the same time, younger respondents were also more likely to approve of the preferential distribution of jobs to Korean nationals. In other words, being more open to migrants and other cultures may be an issue that is separate from the employment issue. Why does the age variable lead to such different outcomes?

Figure 3(a) presents the changes in the unemployment rate by age group from 2010 to 2019. While the overall unemployment rate remains below 4%, the unemployment rate for people in their 20s is far greater. Even though the rate decreased slightly between 2018 and 2019, the unemployment rate for 20 year olds continued to increase. The rate for people in their 30s was better than average, but this group still had the second highest unemployment rate. Figure 3(b) presents the changes in the migrant worker population and marriage-based immigrants from 2013 to 2019. The population of this group has continued to grow. As of 2019, the number of migrant workers was 515,051, and the number of marriage-based immigrants was 173,882.

Taking into account the relatively high unemployment rate among 20 and 30 year olds, along with the influx of migrants, the younger age groups’ approval of preferential job distribution to Korean nationals is quite understandable (Park, 1998; 2000; Sohn, 2019). This younger generation is also highly dissatisfied with the current socio-economic structure of Korean society (Han 2022c). In contrast, other age groups with relatively lower unemployment rates were somewhat free from direct or psychological competition with migrants. From the age variable’s direction of influence on the job debate, we discovered proof of realistic threats. Next, we can analyse another key variable along with age, the region variable. Living in the capital area promoted multicultural receptivity, whereas living in the Honam-Janu region suppressed multicultural receptivity. However, in terms of employment opportunities,
living in the Honam-Jeju region clearly promoted disapproval. Such differences exist in population composition and industry structure. Figure 4(a) provides the population of migrants by region, and Figure 4(b) provides the rate of overseas travel experiences by region.

With regard to the population of migrants by region, their population was highest in the capital area, as 58% of migrant workers and 56% of marriage-based immigrants lived in the capital area. Including foreign students and other foreigners, 60% of all foreign residents lived in the capital area. With regard to the statistics on overseas travel experience, residents of the capital area and Chungcheong region had higher rates of overseas travel experience than residents of other regions. One can interpret this to mean that frequent interactions and encounters with migrants or foreigners, along with foreign cultural experiences via overseas travel, may have increased the multicultural receptivity of residents in those corresponding regions.

Figure 5 provides the elderly population ratio by region. Figure 5(a) reveals that the ratio of elderly to the entire population is highest in the Honam-Jeju region, with the ratio continuing to grow. In the Yeongnam region, the elderly population ratio has grown rapidly, surpassing the elderly population ratio in the Gangwon-Chungcheong region. In contrast, the capital area has a comparatively low elderly population ratio. Since the elderly population ratio is higher in the Honam-Jeju and Yeongnam regions, the multicultural receptivity of the general population may be poor. However, they might have opposed preferential job distribution to Korean nationals because they are less likely to compete with migrants for job opportunities.

Figure 5(b) shows that the ratio of elderly people to the entire population is higher in small/medium-sized cities and rural areas than in metropolitan cities. It is probable that the relatively higher ratio of elderly in rural areas influenced their disapproving views towards preferential job distribution to Korean nationals. Because the elderly and migrant populations in rural areas were unlikely to be in competitive relationships over jobs or resources, the probability for realistic threats was low. What, then, is the relationship between rural areas and multicultural receptivity? Figure 2 revealed that living in rural areas and small cities promoted multicultural receptivity. This growth in multicultural receptivity might be the result of a decreasing rural population and an influx of migrants.

Figure 6(a) indicates that the rural population, in contrast to the urban population, continued to decrease. Figure 6(b) indicates that the ratio of migrants to the total population was higher in small/medium-sized cities and rural areas than in metropolitan cities. Under these circumstances, with a decreasing and aging population, the influx of migrants helped prevent a rapid decline of the population and workforce in rural areas (Kim, 2009a, 2009b; Kim, 2012b; Lee, 2012). International marriages in Korea as of 2019 showed a much greater number of foreign rural areas (Kim, 2009a, 2009b; Kim, 2012; Lee, 2012). As of 2019, over 50% of all migrants lived in the capital area, and only 10% lived in the Honam-Jeju region. With regard to the employment status of foreigners, including migrants in Korea, 44.8% were employed in the mining and manufacturing industries, and 19.4% were employed in service industries. Only 6.7% were employed in the agriculture, forestry, and fishing industries, which is very low compared to other industries.

As Figure 8 indicates, the influx of migrants into rural areas had a positive effect on rural farming communities that were declining due to population; decreases and aging populations; thus, this may work towards promoting receptivity towards multiculturalism (Lee, 2008, 2012; Kim, 2009b; Hiroo and Lee, 2009). At the same time, this also works towards disapproval of preferential job distribution to Korean nationals. With regard to relationships, when respondents had a relationship with a migrant as a family member, relative, or friend, it increased their receptivity towards multiculturalism. Not having a relationship with migrants promoted disapproval of the preferential job distribution to Korean nationals. This is a contradictory finding. When migrants live near a person, this can lead to increased familiarity with them through interactions (Jordan and Duvell, 2003; Bennetts, 2010; Kim, 2010, 2013). As familiarity grows, the person’s understanding of those migrants and their cultures also grows, which should lead to an increased receptivity towards multiculturalism (Kim, 2010). Nevertheless, increased familiarity does not necessarily prevent exclusionary attitudes regarding competition for jobs. Proximity and encounters with migrants promotes multicultural receptivity, but Koreans may still perceive them as rivals for job opportunities. In other words, this ambivalent position may be the result of recognizable realistic threats (McLaren, 2003; Callens et al., 2015; Gorodzeisky and Semyonov, 2019).

One can understand the impact of occupational status in a similar way. In terms of occupational status, working as a temporary/day labourer or in an unpaid family promoted receptivity towards multiculturalism. Since the employment status of many migrants was as low-paying temporary/day labourers, Korean nationals in that occupational status may have more encounters with migrants than Koreans in other occupations (Kim, 2004; Seol, 2012; Lee, 2013; Kim et al., 2014; Kim and Choi, 2018). Daily interactions between these two groups may have contributed to their enhanced receptivity towards multiculturalism (Wessel, 2009; Kim, 2010). However, when it came to job opportunities, they took on an exclusionary stance.

According to Figure 9, which presents the total non-permanent positions in Korea from 2013 to 2019, the number of the positions continued to increase. The percentage of non-permanent workers in the overall work force has also risen, from 32.6% in 2013 to 36.4% in 2019. As shown in Figure 3(b), the average unemployment rate has also grown. Taking these two conditions into account, the job market in Korea is worsening in both quality and in quantity. Under these socio-economic circumstances, a continuous influx of migrants and their employment in non-permanent positions may create conflict between the two groups. This perspective may cause Korean people in these circumstances to approve of preferential job distribution to Korean nationals (Park, 1998, 2000).

To summarise the findings above, intergroup conflicts may manifest as symbolic or realistic threats, depending on the type of relationships
that exist between Korean nationals and migrants. As Korean people meet and come into contact with migrants, they become closer to, and more understanding of, migrants and their cultures, which leads to an increased receptivity towards multiculturalism. On the other hand, when Koreans were not in competitive relationships with migrants, the tension between them decreased. Therefore, contact, rapport, experience, and familiarity with other cultures promoted receptivity, but direct or indirect competition over job opportunities led to realistic threats.

5. Conclusion and implications

The findings of this study have two implications: one from a methodology aspect, and another from the aspect of content about exclusionary attitudes toward migrants. Machine learning algorithms provide a new approach for analysing survey data. This study applied the LightGBM and SHAP algorithms, and analysed the variables in terms of two separate categories: multicultural receptivity and the job debate. In doing so, this study accounted for individual and personal differences in the in-group’s perceptions about the out-group. This combination of two algorithms, LightGBM and SHAP, allowed us to analyse both categorical and continuous data. It also permitted us to consistently analyse which variables strongly influenced the output, and how the value of each variable exerted its influence on the output. By using an intuitively understandable SHAP summary plot, we analysed the importance of each variable, the direction of the variable’s impact on the output, and the differences between variables. Using the algorithmic approaches, this study illuminated new perspectives by presenting the possibility that surveys can be analysed in unconventional ways.

With regard to the deluge of discourse and policies on multiculturalism, Korean society has not voiced a particular opinion. At present, Korean society tries to justify multiculturalism and Koreans passively accept education and policies regarding multicultural receptivity. This study proposes the following questions to Korean society: with regard to migrants’ social status, should we classify them as working class, or another group that is distinct from the working class? Migrants may be working class, considering their socio-economic statuses in the labour market; however, they may be categorised as a separate group that is distinct from the working class, since they do not partake in the mainstream national values required to participate in the society’s system. This is because, unlike foreign brides who remain in Korea and form families, migrant workers, who make up the largest portion of migrants, must return to their home country after a certain period of time. Migrants are an important part of the labour market in Korea, and clearly have complicated relationships with Korean nationals that include both rapport and threats. How should we view these relationships: as a form of symbolic threats that stems from cultural differences, or as a realistic threats over opportunities?

Integrated Threat Theory employed in this study is effective in accounting for individual and personal variations in the in-group’s perceptions towards migrants. Koreans’ stances on the influx of migrants into Korea varied, depending on the type of relationship they had with migrants. Based on those relationships, this study divided threats into symbolic threats (which arise from differences in values and culture) and realistic threats (which arise from competition over resources and opportunities). Direct contact with migrants promoted multicultural receptivity, based on an understanding of their culture. However, having competitive relationships with migrants over resources and opportunities led to realistic threats. For example, the younger generations are more open to multiculturalism, but they are more competitive with migrants due to their socio-economic conditions. The older generations are more exclusive to multiculturalism, but their socio-economic conditions make them less competitive with migrants.

Before the influx of migrant workers and foreign brides, extreme exclusion characterized the history of ethnic minorities in Korea. The mechanisms of exclusion were extreme discrimination and disdain towards Chinese people and interracial children, followed by expulsion from ordinary society. Chinese people who resided in Korea were a symbolic group with extremely limited civil rights, such as the freedom to choose their occupation and the right to own property. The interracial children provided another extreme example that reminded us of Western society’s exclusion of insane people during the early Modern Age. Similar to insane people in the 17th century, as mentioned in Foucault (1989), these people were expelled from people’s consciousness.

Through encounters with migrants, education, and experiences with other cultures, multicultural receptivity in Korean society has improved greatly. However, Koreans’ exclusionary attitudes toward migrants will increase greatly, irrespective of their level of receptivity, when work issues amplify the realistic threats. The immigration phenomenon did not draw people’s attention until low-income people, who suffered a decline in their living standards during a recession, began to feel concerned and perceived migrants as rivals. If one adds nationalistic discourse to this public sentiment, hatred toward foreigners, which was banned and limited to a certain group, becomes prevalent among the public. Korea is not free from COVID-19 Pandemic that broke out in 2020. It has been about two years since the outbreak of COVID-19, and in the year 2022, recession, bankruptcy, high unemployment and, in particular, mass layoffs, have become a reality. As Koreans lose employment opportunities and economic polarization grows, what will Korean people think of migrants? Will they embrace migrants or exclude them harshly, as they did before?

This study proposed a new analytical approach that combined survey data and machine learning algorithms. By applying this approach, this study provided a new perspective on exclusionary attitudes toward migrants. As for the variables that this study did not discuss fully, follow-up studies should investigate them further.

Declarations

Author contribution statement

Seungwoo Han:Concepted and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data associated with this study has been deposited at https://mdis.kostat.go.kr/eng/pageLink.do?link=mdisDataService.

Declaration of interest’s statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2022.e10087.

References

Baumeister, R.F., Vohs, K.D., 2007. Realistic group conflict theory. Encycl. Soc. Psychol. 2, 725–726.
Beilanger, D., Lee, H.K., Wang, H.Z., 2010. Ethnic diversity and statistics in East Asia: foreign brides’ surveys in Taiwan and South Korea. Ethn. Racial Stud. 33, 1108-1135.
Bennett, C., 2010. The impact of transformational learning on individuals, families and communities. Int. J. Lifelong Educ. 22 (5), 457-480.
Bonaccorsi, G., 2018. Mastering Machine Learning Algorithms: Expert Techniques to Implement Popular Machine Learning Algorithms and fine-tune Your Models. Packt Publishing Ltd, Birmingham.
Brunori, P., Palmanio, F., Peragine, V., 2019. Inequality of opportunity in sub-Saharan Africa. Appl. Econ. 41 (12), 1628-6458.
Brunori, P., Neidhoffer, G. 2020. The evolution of inequality of opportunity in Germany: a machine learning approach. ECIQEIN Working Paper Series 2020-514.
Buehler, M., Fabbe, K.E., Han, K.J., 2020. Community-level postmaterialism and anti-immigrant attitudes: original survey on sub-saharan African migrants in the Middle East. Int. Stud. Q. 64, 669-683.
Burrell, J., 2016. How the machine ‘thinks’: understanding opacity in machine learning algorithms. Big Data & Society 3 (1), 1–12.
Campbell, D.T., 1965. Understanding female marriage immigrants to rural areas in South Korea: the making of a multicultural society. Ethn. Racial Stud. 32 (1), 70–92.
Kim, A.E., 2009a. Global migration and South Korea: foreign workers, foreign brides and the making of a multicultural society. Ethn. Racial Stud. 32 (1), 70–92.
Kim, A.E., 2009b. Demography, migration and multiculturalism in South Korea. Asia-Pac. J. 7 (6), 1–19.
Kim, B.T., Choi, G.G., 2018. Understanding female marriage immigrants’ participation in economic activities and their working conditions in South Korea. Asian Pac. Migr. J. 27 (3), 343–367.
Kim, S., 2012a. Racism in the global era: analysis of Korean media discourse around migrants, 1990-2009. Discourse Soc. 23 (6), 657-678.
Kim, H., Lee, S.Y., Choi, J.H., 2014. Employment and poverty status of female marriage immigrants in South Korea. Asian Pac. Migr. J. 23 (2), 129–154.
Kim, H.K., 2012b. Marriage migration between South Korea and Vietnam: a gender perspective. Asian Perspect. 36 (3), 531–563.
Kim, Y., 2008. Changing an agent-based learning under the influence of migration: South Korea. Int. J. Lifelong Educ. 29 (2), 255-272.
Kim, J.K., 2011. The politics of culture in multinational in Korea. J. Ethn. Migr. Stud. 37 (10), 1583–1604.
Kim, S., 2013. Citizenship projects for marriage migrants in South Korea: intersecting motherhood with ethnicity and class. Soc. Polit. Int. Stud. Gend. State Soc. 20 (4), 455–481.
Kim, S., 2015. Soft talk, hard realities: multiculturalism as the South Korean government’s decoupled responses to international migration. Asian Pac. Migr. J. 24 (2), 51–74.
Knowles, V., 2016. Strangers at Out Gates: Canadian Immigration and Immigration Policy, vol. 1540–2015 (fourth ed.), Dundurn Toronto, Toronto, Canada.
Kwak, Y., Joo, Y.S., Woo, S.H., 2019. Social Inclusion of Immigrants: Current State and Policy Implications. Korea Institute for Health and Social Affairs, Sejong, South Korea.
Kwon, S., 2009. A study on the receptivity of local government officials regarding multi-cultural society. 21st Century Polit. Sci. Rev. 19 (1), 353–378.
Lefranc, A., Kundo, T. 2020. Inequality of Opportunity in India Society. Hal-02539364. Les Éditions de l’École Polytechnique.
Lee, H., 2012. The political economy of cross-border marriage: economic development and social reproduction in Korea. Fem. Econ. 18 (2), 177–200.
Lee, H.J., 2010. Anti-bias education in elementary school in multi-cultural society. J. Ethics Educ. Stud. 22, 253–272.
Lee, Y., 2009. Migration, migrants, and contested ethno-nationalism in Korea. Crit. Asian Stud. 41 (3), 363–380.
Lee, H.K., 2008. International marriage and the state in South Korea: focusing on governmental policy. Citizen. Stud. 12 (1), 107–123.
Lee, H.K., 2013. Employment and life satisfaction among female marriage migrants in South Korea. Asian Pac. Migr. J. 22 (2), 195–219.
Lundberg, S.M., Lee, S.L., 2017a. Interpretation and prioritization. Proceedings of the 31st Conference on Neural Information Processing Systems, 4–9 December 2017, Long Beach, CA, USA.
Lundberg, S.M., Lee, S.L., 2017b. Consistent feature attribution for tree ensembles. Proceedings of the 2017 ICML Workshop on Human Interpretability in Machine Learning (WHI2017), 10 August 2017, Sydney, NSW, Australia.
Maloney, K.O., Schmid, M., Weller, D.E., 2012. Applying additive modeling and gradient boosting to assess the effects of watershed and reach characteristics on riverine assemblages. Methods Ecol. Evol. 3, 116–128.
McGhee, D., 2005. Intolerant Britain? Hate, Citizenship and Difference. Open University Press, New York.
McLaren, L.M., 2001. Immigration and the new politics of inclusion and exclusion in the European Union: the role of elites and the EU on individual-level opinions regarding European and non-European immigrants. Eur. J. Polit. Res. 38, 81–108.
McLaren, L.M., 2003. Anti-Immigrant prejudice in Europe: contact, threat perception, and support for the exclusion of immigrants. Soc. Forces 81 (3), 909–939.
Meng, Q., Ke, G., Wang, T., Chen, W., Ye, Q., Ma, A.M., Liu, Y.T., 2016. A communication-efficient parallel algorithm for decision trees. 30th conference on neural information processing systems. Adv. Neural Inf. Process. Syst. 29, 1279–1287.
Min, P.G., 1991. Cultural and economic boundaries of Korean ethnicity: a comparative analysis. Ethn. Racial Stud. 14 (2), 225–241.
Mulainathan, S., Spiess, J., 2017. Mach. Learn./Appl. Econ. Approach 31 (2), 87–106.
Murphy, K.P., 2012. Machine Learning: A Probabilistic Perspective. The MIT Press, Cambridge.
Murthy, S.K., 1998. Automatic construction of decision trees from data: a multidisciplinary survey. Data Min. Knowl. Discov. 2 (4), 345–389.
Nelson, S.M., 2010. Politics of ethnicity in prehistoric Korea. In: Prexcel, R.W., Mozokowski, S.A. (Eds.), Contemporary Archaeology in Theory: the New Pragmatism. Blackwell Publishing Ltd, Oxford, pp. 290–300.
Oh, J., Yun, K., Miao, U., Kim, T.S., Chae, J.H., 2019. Identifying depression in the national Health and Nutrition Examination Survey data using a deep learning classification-algorithm evaluation. J. Affect. Disord. 257, 623–631.
Hogg, M.A., 2013. Intergroup relations. In: Delamater, J., Amanda, W. (Eds.), Handbooks of Sociology and Social Research. Handbook of Social Psychology (second ed.). Springer, New York, pp. 479-501.
Hosken, P., 2009. A study of teachers’ perception of the children of migrants: a case study of one primary school in Seoul, South Korea. J. Multicult. Soc. Hist. 12 (1), 263–295.
Jordan, B., Duvall, F., 2015. Migration: The Boundaries of Equality and Justice. Polity, London.
Kaya, Y., 2012. Civilizing vs destructive globalization? A multi-level analysis of anti-immigrant prejudice in Japan. Int. Comp. Soc. 53 (3), 23–44.
Ke, G., Meng, Q., Tie, Y., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, Y.T., 2017. LightGBM: a highly efficient gradient boosting decision tree. Adv. Neural Inf. Process. Syst. 30, 3146–3154.
Kim, W.B., 2004. Migration of foreign workers into South Korea: from periphery to semi-periphery in the global labor market. Asian Surv. 44 (2), 316–335.
Park, Y.B., 2000. Unskilled foreign labor and Korea’s economic crisis. Asian Pac. Migrat. J. 9 (4), 483–506.
Powell, K., 2017. Brexit positions: neoliberalism, austerity and immigration—the (im) possibilities? Of political revolution. Dialect. Anthropol. 41, 225–240.
Powers, D.M.W., 2011. Evaluation: from prevision, recall and F-measure to ROC informedness, markedness & correlation. J. Mach. Learn. Technol. 2 (1), 37–63.
Rzepnikowska, A., 2019. Racism and xenophobia experienced by Polish migrants in the UK before and after Brexit vote. J. Ethnic Migrat. Stud. 45 (1), 61–77.
Rudra, N., 2014. Welfare states in developing countries: unique or universal? J. Polit. 69, 378–396.
Scheve, K.F., Slaughter, M.J., 2006. Labor market competition individual preferences over immigration policy. Rev. Econ. Stat. 83 (1), 133–145.
Schmidt, P., Quandt, M., 2018. National identity, nationalism, and attitudes towards migrants in comparative perspective. Int. J. Comp. Social. 59 (5-6), 355–361.
Seol, D.H., 2012. The citizenship of foreign workers in South Korea. Citizen. Stud. 16 (1), 119–133.
Shapley, L.S., 1953. A value for n-person games. In: Tucker, A.W., Kuhn, H.W. (Eds.), Contributions to the Theory of Games, vol I. Princeton University Press, Princeton, pp. 1-17.
Shin, G.W. 2006. Ethnic Nationalism in Korea. Stanford University Press, Standford.
Sohn, I., 2019. The contentious politics of youth unemployment: comparing Korean and Taiwanese. Kor. J. Int. Stud. 17 (1), 55–77.
Song, M.Y., 2010. The factors influencing multicultural receptivity in Korea. Minjok Yeonku 41, 74–97.
Stangor, C.L., Sullivan, A., Ford, T.E., 1991. Affective and cognitive determinants of prejudice. Soc. Cognit. 9 (4), 359–380.
Stephan, W.G., Stephan, C.W., 1996. Predicting prejudice. Int. J. Intercult. Relat. 20 (3-4), 409–426.
Stephan, W.G., Stephan, C.W., Gudykunst, W.B., 1999a. Anxiety in intergroup relations: a comparison of anxiety/uncertainty management theory and integrative threat theory. Int. J. Intercult. Relat. 23 (4), 613–628.
Stephan, W.G., Ybarra, O., Bachman, G., 1999b. Prejudice toward immigrants. J. Appl. Soc. Psychol. 29 (11), 2221–2237.
Taifel, H. 1981. Social stereotypes and social groups. In: Turner, J.C., Giles, H. (Eds.), Intergroup Behavior. University of Chicago Press, Chicago.
Taifel, H., Turner, J.C., 1986. The social identity theory of intergroup behavior. In: Worchel, S., Austin, W.G. (Eds.), Psychology of Intergroup Relation. Nelson-Hall Chicago, pp. 7-24.
Wessel, T., 2009. Does diversity in urban space enhance intergroup contact and tolerance? Geogr. Ann. B Hum. Geogr. 91 (1), 5-17.
Won, S.Y., 2011. Perception toward foreigners and state policy for foreigners: focusing on citizen in Seoul metropolitan city. Publ. Policy Rev. 25 (4), 27–55.
Zarate, M.A., Garcia, B., Garza, A.A., Hitlan, R.T., 2004. Cultural threat and perceived realistic group conflict as dual predictors of prejudice. J. Exp. Soc. Psychol. 40 (1), 99–105.