Exploring Factors Associated with the Social Discrimination Experience of Children from Multicultural Families in South Korea by using Stacking with Non-linear Algorithm

Haewon Byeon
Department of Medical Big Data, College of AI Convergence
Inje University, Gimhae 50834, Republic of Korea

***Abstract*—The number of children from multicultural families is increasing rapidly along with quickly increasing multicultural families. However, there are not enough surveys and basic researches for understanding the characteristics of multicultural children and issues such as social discrimination. This study discovered the machine learning model with the best performance for predicting the social discrimination experience of children from multicultural families by comparing the prediction performance (accuracy) of individual prediction models and stacking ensemble models. This study analyzed 19,431 adolescents (between 19 and 24 years old: 9,835 males and 9,596 females) among the children of marriage immigrants. This study used random forest (RF), rotation forest, artificial neural network (ANN), and support vector machine (SVM) for the base model. Logistic regression algorithm was applied for the meta model. Each machine learning model was built through 5-fold cross-validation. Root-mean-square-error (RMSE), index of agreement (IA), and variance of errors (Ev) were used to evaluate the prediction performance of the developed models. The results of this study indicated that the prediction performance of the rotation forest-logistic regression model had the best performance. The future studies need to explore stacking ensemble models with the best performance through combining a base model and a meta model by using various machine learning algorithms such as clustering and boosting.**

***Keywords*—Stacking ensemble; meta model; root-mean-square-error; index of agreement; rotation forest**

I. Introduction

The number of foreigners residing in South Korea exceeded 2 million as of 2019. This accounts for 3.69% of the South Korean population, which is not a high percentage. However, it is recognized as a noticeable phenomenon in Korean society because the number of immigrants has increased rapidly over the past decade and immigrants are easily distinguishable due to differences in appearance and language [1]. In particular, as this issue has become linked to the marriage to men living in rural areas or men with low-income in urban areas since 2002, the number of multicultural families has reached 900,000 as of 2016 [2]. The number of immigrants will increase more as the population of South Korea will decrease due to the aging and low birth rate [3]. It has drawn more attention because the population composition will be diversified further due to this [3].

The multicultural family means a family made by uniting people with different nationalities or races through international marriage and other methods. South Korea prepared the "Measures to Support the Social Integration for Female Marriage Immigrant Families, Multi-racial People, and Immigrants" in 2006 to help multicultural families settle in South Korea stably. As the Multicultural Families Support Act was enacted in 2008, she strengthened the support for multicultural families at the policy level. As a result, social security and legal status were guaranteed for marriage immigrants.

As the number of foreigners residing in South Korea rapidly increases, the number of children from multicultural families (e.g., international marriage families and foreign workers' families) is also increasing. Furthermore, as they attend schools, the possibility of conflict due to cultural differences has increased according to the increased personal and cultural contacts.

Nevertheless, in South Korea, social policies for multicultural families have mainly focused on employment or welfare for marriage immigrants and foreign workers [4,5]. Moreover, previous studies on multicultural families [4,6] have been conducted to examine only limited individual aspects such as socioeconomic characteristics, welfare level, human rights discrimination, employment status, and policy analysis. However, there are still not enough studies on the overall social discrimination experiences of children from multicultural families. Children from multicultural families (international marriage families) can be divided into children born in South Korea and those who have entered South Korea after being born in other countries. Since it has been reported that children could not adapt to South Korea well due to the unique characteristics of multicultural families and various changes that they experience during adolescence [7], it is necessary to expand the social foundation that can help them adapt to South Korea well for social integration.

In summary, the number of children from multicultural families is increasing rapidly along with quickly increasing multicultural families. However, there are not enough surveys and basic researches for understanding the characteristics of multicultural children and issues such as social discrimination. Therefore, it is needed to identify the characteristics of

multicultural children and seek new policies that reflect them to prepare policies that encompass various problems including the social adaptation of multicultural children in preparation for a rapidly changing multicultural society.

Previous studies [8,9,10,11,12] on the adolescents from multicultural families in South Korea reported the difficulties in peer relations, social support, family support, and language as factors related to discrimination experiences. Most of them used regression analysis for a prediction algorithm. Regression analysis is efficient for detecting individual risk factors, but it is limited in identifying multiple risk factors [13,14]. As a way to overcome this limitation, recent social science studies [15,16] have used predictive modeling based on big data-based machine learning. However, since these prediction studies are based on individual prediction algorithms, the bias existing in each algorithm may be reflected in the prediction results.

This study identified the predictors of social discrimination experiences of children from multicultural families in South Korea by using individual prediction models based on machine learning and reduced the potential bias risk of the models by combining them into a stacking ensemble learning model. Moreover, this study discovered the machine learning model with the best performance for predicting the social discrimination experience of children from multicultural families by comparing the prediction performance (accuracy) of individual prediction models and stacking ensemble models.

II. Methods and Materials

A. Data Source

The data source of this study was the "Study on the National Survey of Multicultural Families [17]" in 2012, which was jointly surveyed by the Ministry of Health, Welfare and Family Affairs, the Ministry of Justice, and the Ministry of Gender Equality and evaluated multicultural families residing in South Korea. The Study on the National Survey of Multicultural Families was conducted to develop policies customized for multicultural families by identifying their living conditions and welfare needs. The survey items consisted of the general characteristics, employment, economic level, marriage, health, and health care of multicultural families. The survey targets for the national survey of multicultural families were 154,333 families, all marriage immigrants. In addition to marriage immigrants, this survey also evaluated the spouses and children of marriage immigrants separately. The target subjects were sampled based on the status of alien residents living in 16 cities and provinces and the basic status of multicultural families collected by the Ministry of Public Administration and Security. Since this survey collected data from all target samples, a sample design was not needed and the survey was conducted from July 20 to October 31, 2012. The multicultural families used for this study were (1) families composed of marriage immigrants and South Korean who obtained South Korean nationality by birth, acknowledgment, or naturalization and (2) families composed of foreigners who obtained South Korean nationality by acknowledgment or naturalization and South Koreans who obtained South Korean nationality by birth, acknowledgment, or naturalization in accordance with the Multicultural Families Support Act. This study analyzed 19,431 adolescents (between 19 and 24 years old: 9,835 males and 9,596 females) among the children of marriage immigrants.

B. Measurements and Definitions of Variables

The target variable (label) was defined as social discrimination experience (yes or no). Features were gender, age, residence (countryside or city), highest level of education (elementary school graduation and below, middle school graduation, high school graduation, or college graduation or higher), Korean reading level (good, average, or poor), Korean speaking level (good, average, or poor), Korean writing level (good, average, or poor), Korean listening level (good, average, or poor), learning support experience (yes or no), economic activity (yes or no), the experience of using a support center for multi-cultural families (yes or no), learning Korean (yes or no), Korean society adaptation training (yes or no), career counseling (yes or no), and social welfare center use (yes or no).

C. Single Machine Learning Algorithm (base model): Support Vector Machine (SVM)

SVM is a machine learning algorithm that finds an optimal decision boundary, which is a linear separation that optimally separates a hyperplane by transforming training data into a high dimension through nonlinear mapping [18]. For example, when A=[a,d] and B=[b,c] are non-linearly separable in 2D, it becomes linearly separable when they are mapped in 3D. Therefore, if an appropriate nonlinear mapping is applied to a sufficiently large dimension, data with two classes can always be separated in a hyperplane. The concept of SVM is presented in Fig. 1.

D. Random Forest

Random forest is an algorithm that randomly learns a number of decision trees. It uses a number of bootstrap samples. After generating a decision tree for each sample, the output value is predicted using the decision tree most frequently used among the generated decision tree when new data is input [20]. The concept of random forest is presented in Fig. 2.

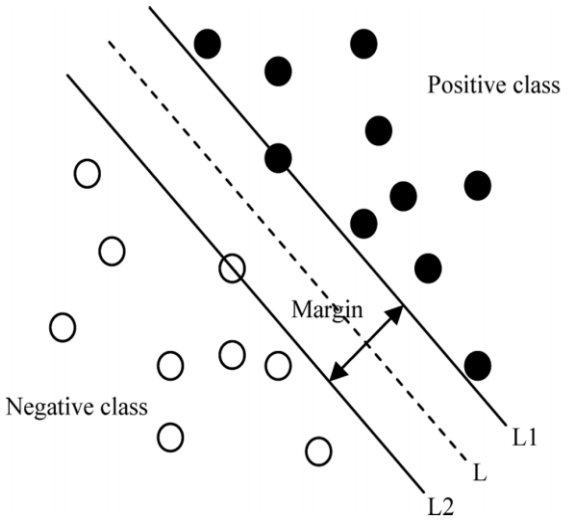

Fig. 1. Concept of SVM [19].

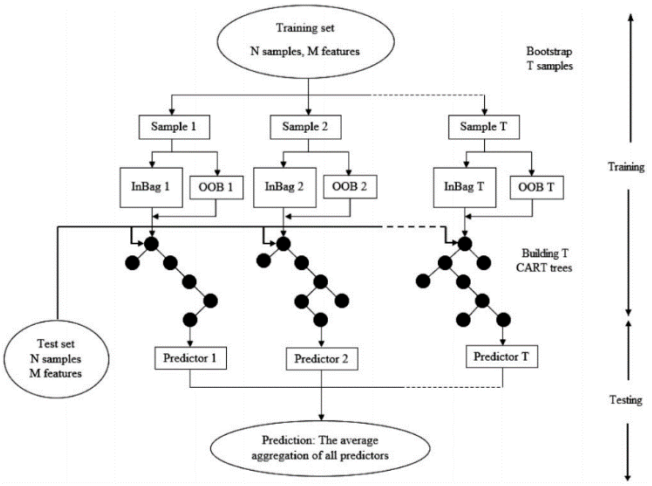

Fig. 2. Structure of Random Forest Algorithm [21].

E. Rotation Forest

Rotation forest is one of the random forest algorithms that performs learning while rotating the data axis by applying principal component analysis (PCA) to the training data. Rotation forest generates classifier ensembles based on feature extraction after excluding random features from the previous feature set used for learning. Principal component analysis (PCA) is performed on randomly divided subsets and training is conducted by rotating the data dimension [22]. Through this process, robust characteristics can be obtained for the input data showing complex distribution [23]. The performance procedure of the rotation forest is presented in Fig. 3.

F. ANN (Artificial Neural Network)

ANN is an algorithm created based on the neural network structure of the human brain. It is composed of an input layer that inputs the target data, a hidden layer (or hidden layers) that is an intermediate step, and an output layer that shows the result. Every layer consists of a number of nodes, and only information that exceeds the threshold is passed to the next layer through the activation function. It is possible to predict the result in the output layer after deriving only the necessary information through this. The concept of ANN is presented in Fig. 4.

G. Stacking Ensemble (Meta Model)

This study predicted social discrimination experiences by using stacking ensemble techniques. Stacking ensemble techniques are superior in generalization and robustness to single machine learning models and have been used for classification and prediction in various fields [26,27,28,29]. It is a method of creating a new model by combining different models as if stacking them [30]. It improves the performance of the final model by taking advantage of each model and supplementing its weaknesses while going through the two stages (base and meta) [30].

This study used random forest (RF), rotation forest, ANN, and SVM for the base model. Logistic regression algorithm was applied for the meta model. The regression algorithm is the simplest method to increase the reliability of the base model while maximizing the generality and stability of the model. Feng et al., (2020) [31] reported that it would overfit the training data less probably. Due to this reason, the regression algorithm has been used as a meta model of the stacking ensemble algorithm in many recent publications [31,32], and this study also used it as a meta model for the same reason. The structure of the finally constructed stacking ensemble model is presented in Fig. 5.

Training Phase

Given

- X: the objects in the training data set (an $N \times n$ matrix)
- Y: the labels of the training set (an $N \times 1$ matrix)
- L: the number of classifiers in the ensemble
- K: the number of subsets
- $\{\omega_1, \ldots, \omega_c\}$: the set of class labels

For $i = 1 \ldots L$

- Prepare the rotation matrix R_i^a:
  - Split $\mathbf{F}$ (the feature set) into K subsets: $\mathbf{F}_{i,j}$ (for $j = 1 \ldots K$)
  - For $j = 1 \ldots K$
    - Let $X_{i,j}$ be the data set X for the features in $\mathbf{F}_{i,j}$
    - Eliminate from $X_{i,j}$ a random subset of classes
    - Select a bootstrap sample from $X_{i,j}$ of size 75% of the number of objects in $X_{i,j}$. Denote the new set by $X'_{i,j}$
    - Apply PCA on $X'_{i,j}$ to obtain the coefficients in a matrix $C_{i,j}$
  - Arrange the $C_{i,j}$, for $j = 1 \ldots K$ in a rotation matrix R_i as in equation (1)
  - Construct R_i^a by rearranging the the columns of R_i so as to match the order of features in $\mathbf{F}$.
- Build classifier D_i using (XR_i^a, Y) as the training set

Classification Phase

- For a given $\mathbf{x}$, let $d_{i,j}(\mathbf{x}R_i^a)$ be the probability assigned by the classifier D_i to the hypothesis that $\mathbf{x}$ comes from class ω_j. Calculate the confidence for each class, ω_j, by the average combination method:

$$\mu_j(\mathbf{x}) = \frac{1}{L}\sum_{i=1}^{L} d_{i,j}(\mathbf{x}R_i^a), \quad j = 1, \ldots, c.$$

- Assign $\mathbf{x}$ to the class with the largest confidence.

Fig. 3. Procedure of Rotation Forest [24].

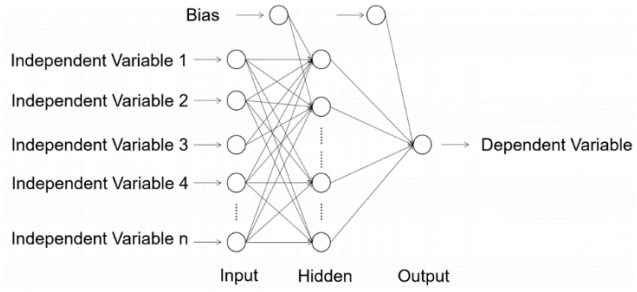

Fig. 4. Algorithmic Structure of a Typical ANN [25].

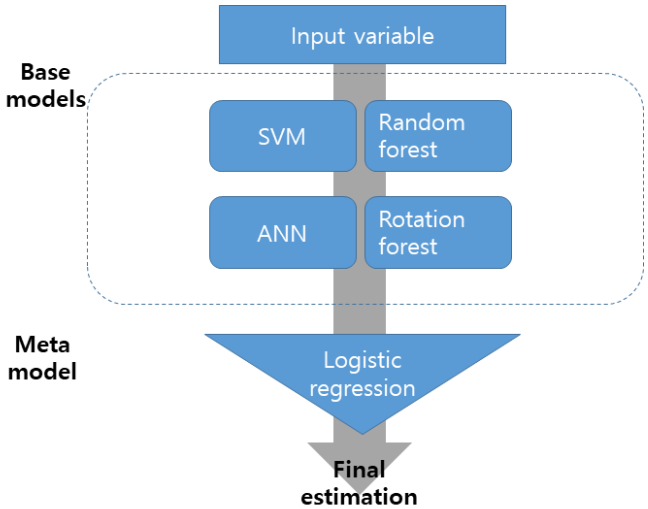

Fig. 5. The Structure of the Stacking Ensemble.

H. Validation of the Models

Each machine learning model was built through 5-fold cross-validation. This method validates the validity of learning by randomly dividing the entire sample into equal-sized five groups, validating it by using one of the groups as a validation dataset and the other groups as training datasets, and repeating this procedure five times. Root-mean-square-error (RMSE), index of agreement (IA), and variance of errors (Ev) were used to evaluate the prediction performance of the developed models. A lower RMSE indicates the higher accuracy of a prediction model. When IA is closer to 1 and Ev is lower, a model is more stable.

III. Results

Table I shows the general characteristics of adolescents from multicultural families in South Korea according to the presence of social discrimination experience. Among the all subjects (19,431 adolescents), 15.6% (3,035 adolescents) experienced social discrimination. The result of chi-square test revealed that residence, gender, highest level of education, the experience of using a support center for multi-cultural families, Korean speaking level, Korean listening level, Korean reading level, Korean writing level, career counseling, learning Korean, and Korean society adaptation training were significantly ($p<0.05$) different between adolescents from multicultural families with social discrimination experience and those without social discrimination experience.

TABLE I. General Characteristics of Adolescents from Multicultural Families in South Korea, n (%)

Variables	Social discrimination experience		p
	Yes (n=3,035)	No (n=16,396)	
Residence			<0.001
City	2,659 (16.2)	13,802 (83.8)	
Countryside	375 (12.6)	2,594 (87.4)	
Highest level of education			
Elementary school graduation and below	26 (12.9)	176 (87.1)	
Middle school graduation	394 (15.9)	2,084 (84.1)	
High school graduation	2,164 (15.3)	12,024 (84.7)	
College graduation or higher	451 (17.6)	2,112 (82.4)	
Gender			<0.001
Male	1,898 (19.3)	7,938 (80.7)	
Female	1,137 (11.8)	8,459 (88.2)	
Korean speaking level			<0.001
Good	2,176 (14.4)	12,939 (85.6)	
Average	635 (21.4)	2,330 (78.6)	
Poor	224 (16.6)	1,128 (83.4)	
Korean reading level			<0.001
Good	2,117 (14.0)	13,016 (86.0)	
Average	496 (19.6)	2,033 (80.4)	
Poor	422 (23.8)	1,348 (76.2)	
Korean writing level			<0.001
Good	2,033 (13.8)	12,731 (86.2)	
Average	506 (18.7)	2,196 (81.3)	
Poor	496 (25.2)	1,470 (74.8)	
Korean listening level			<0.001
Good	2,220 (14.4)	13,165 (85.6)	
Average	677 (24.1)	2,127 (75.9)	
Poor	137 (11.0)	1,105 (89.0)	
Economic activity			0.659
No	1,790 (15.7)	9,598 (84.3)	
Yes	1,245 (15.5)	6,798 (84.5)	
Korean society adaptation training			
No	2,888 (15.2)	16,104 (84.8)	
Yes	146 (33.3)	292 (66.7)	
Experience of career counseling			<0.001
No	2,611 (14.3)	15,602 (85.7)	
Yes	424 (34.8)	794 (65.2)	
Learning support experience			<0.001
No	2,526 (14.1)	15,336 (85.9)	
Yes	509 (32.4)	1,060 (67.6)	
Experience of using a support center for multi-cultural families			<0.001
Do not even know that such center exists	1,692 (18.8)	7,332 (81.3)	
Know the center but never used it before	1,125 (12.1)	8,198 (87.9)	
Not only know the center but also have used it before	218 (20.1)	867 (79.9)	

The prediction performance (i.e., RMSE, IA, and Ev) of the eight machine learning models for predicting social discrimination experience is presented in Fig. 6, 7, and 8, respectively. The analysis results of this study indicated that the prediction performance of the rotation Forst-logit regression model (RMSE = 0.15, IA = 0.72, and Ev =0.41) had the best performance.

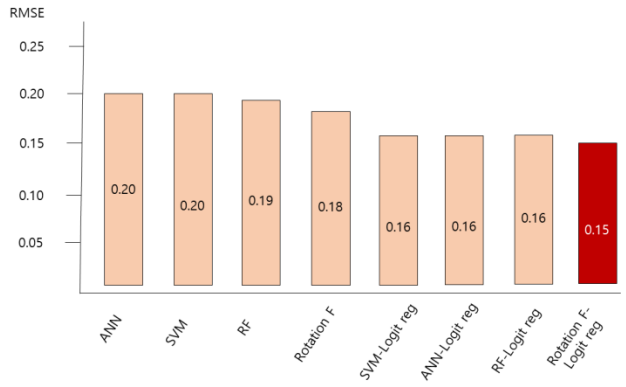

Fig. 6. RMSE Comparison of Machine Learning Models for Predicting Social Discrimination.

ANN=Artificial neural network; SVM=Support Vector Machine, RF=Random forest; Rotation F=Rotation forest; SVM-Logit reg=SVM-Logistic regression; ANN-Logit reg=Artificial neural network-Logistic regression; RF-Logit reg=Random forest-Logistic regression; Rotation F-Logit reg=Rotation forest-Logistic regression.

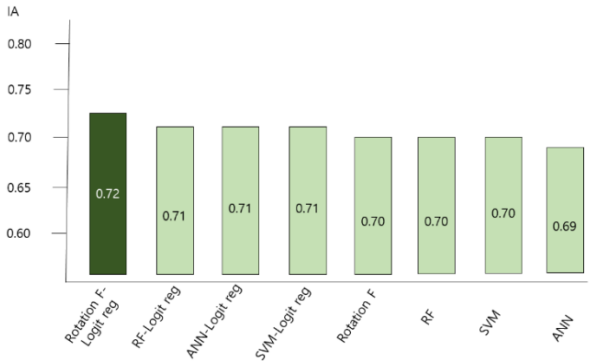

Fig. 7. IA Comparison of Machine Learning Models for Predicting Social Discrimination.

ANN=Artificial neural network; SVM=Support Vector Machine, RF=Random forest; Rotation F=Rotation forest; SVM-Logit reg=SVM-Logistic regression; ANN-Logit reg=Artificial neural network-Logistic regression; RF-Logit reg=Random forest-Logistic regression; Rotation F-Logit reg=Rotation forest-Logistic regression.

The normalized importance of each variable of the rotation forest-logit regression model is presented in Fig. 9. The model confirmed that Korean society adaptation training, learning Korean, gender, the experience of using a multicultural family support center, and career counseling were major variables with high weight in the social discrimination experience of children from multicultural families in South Korea. Especially, Korean society adaptation training was the most important factor in the final model.

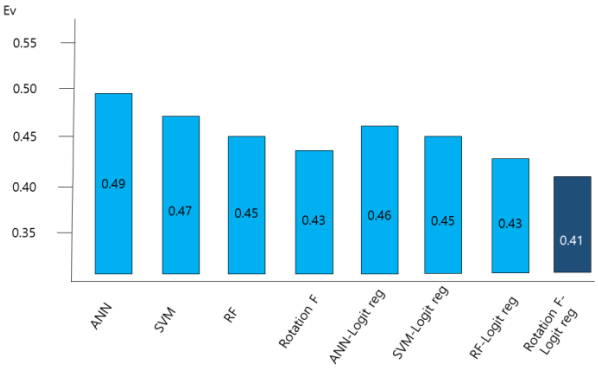

Fig. 8. Ev Comparison of Machine Learning Models for Predicting Social Discrimination.

ANN=Artificial neural network; SVM=Support Vector Machine, RF=Random forest; Rotation F=Rotation forest; SVM-Logit reg=SVM-Logistic regression; ANN-Logit reg=Artificial neural network-Logistic regression; RF-Logit reg=Random forest-Logistic regression; Rotation F-Logit reg=Rotation forest-Logistic regression.

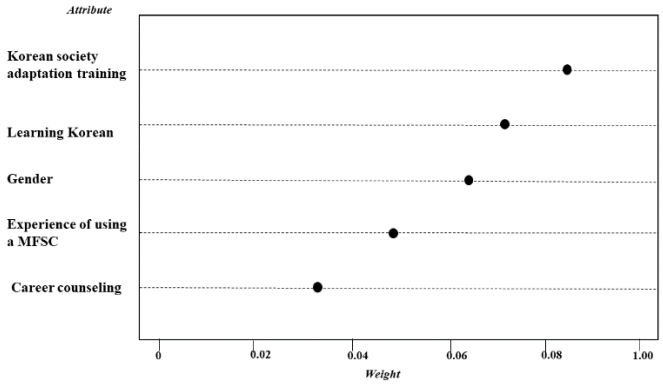

Fig. 9. The Importance of Variables in the Prediction Model for Discrimination Experience of Children from Multicultural Families in South Korea (only the top 5 Variables are Presented).

MFSC=multicultural family support center

IV. Conclusion

This study compared the accuracy of models for predicting the social discrimination experience of children from multicultural families in South Korea by using eight machine learning algorithms, and confirmed that the rotation forest-logit regression model based on the stacking ensemble algorithm had the best prediction performance. In particular, the prediction model based on the stacking ensemble had improved accuracy (RMSE = 0.04-0.05) than other models and more stable (IA= 0.02-0.03) than other models. The results of this study support the possibility that the meta-model's prediction performance can be superior to the single prediction model for not only unstructured data such as videos and images but also structured data such as social science data. However, Lee & Kim (2020) [33] also reported that stacking ensemble algorithms had a longer execution time (runtime) than single machine learning algorithms, a limitation. Therefore, future studies using stacking ensembles need to evaluate the prediction performance comprehensively by comparing execution time (runtime) as well as accuracy. It is also needed to explore stacking ensemble models with the best performance through combining a base model and a meta model by using various machine learning algorithms such as clustering and boosting.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1B07041091, NRF-2019S1A5A8034211, NRF-2021S1A5A8062526).

References

[1] B. G. Koo, Multiculturalism and transnational migrants: a case study of Wongok-dong, an immigrant-dominated area of Ansan city in South Korea. Cross - Cultural Studies, vol. 19, no. 2, pp. 5-51, 2013.

[2] B. J. Park, The role of local government for social Integration with multicultural family. Journal of North-east Asian Cultures, vol. 51, p. 285-307, 2017.

[3] E. H. Chae, Analysis of trends of 'an investigation on multicultural families in Korea' at the Korean Statistical Information Service(KOSIS), vol. 8, no. 11, pp. 11-20, 2018.

[4] J. H. Shin, A study on comparative home environmental factor effect of delinquency in multi-cultural youth. Journal of Korean Public Police and Security Studies, vol. 11, no.2, pp.1-20, 2014.

[5] J. Kim, Problem about policies for multi-cultural society and social integration. Low and Social Study, vol. 11, no. 2, pp. 349-368, 2011.

[6] K. S. Ahn, A plan of multi-cultural adolescent`s healthy upbringing. Korean Journal of Youth Studies, vol. 16, no. 7, pp. 99-126, 2009.

[7] M. Tienda, and R. Haskins, Immigrant children: introducing the issue. The Future of Children, vol. 21, no. 1, pp. 3-18, 2011.

[8] H. M. Kim, W. J. Seo, and S. H. Choi, Experiences of discrimination and psychological distress of children from multicultural families: examining the mediating effect of social support. Korean Journal of Social Welfare Studies, vol. 42, no. 1, pp. 117-149, 2011.

[9] S. Cha, and B. Hyeon, A systematic review on factors influencing multicultural acceptance in Korean adolescents. Journal of the Korea Academia-Industrial cooperation Society, vol. 19, no. 7, pp. 207-213, 2018.

[10] S. J. Kim, A study on the influence of experiences of discrimination to multicultural families' adolescents on their characteristics - focused on raw-data of National Survey of Multicultural Families 2012. The Journal of Asiatic Studies, vol. 58, no. 3, pp. 6-41, 2015.

[11] Y. S. Choi, Personal characteristics, ethnic identity, experience of discrimination, self-esteem, and problem behavior of Korean-Japanese multicultural adolescents. Korean Journal of Family Welfare, vol. 17, no. 2, pp. 49-71, 2012.

[12] A. R. Lee, J. Lee, and B. Y. Son, A qualitative study on the multi-cultural adolescents experience of career barrier. Korean Journal of Youth Studies, vol. 25, no. 11, p. 35-64, 2018.

[13] H. Byeon, Predicting the anxiety of patients with Alzheimer's dementia using boosting algorithm and data-level approach, International Journal of Advanced Computer Science and Applications, vol. 12, no. 3, pp. 107-113, 2021.

[14] H. Byeon, Comparing ensemble-based machine learning classifiers developed for distinguishing hypokinetic dysarthria from presbyphonia. Applied Sciences, vol. 11, no. 5, pp. 2235, 2021.

[15] H. Byeon, S. Cha, and K. Lim, Exploring factors associated with voucher program for speech language therapy for the preschoolers of parents with communication disorder using weighted random forests. International Journal of Advanced Computer Science and Applications, vol. 10, no. 5, pp. 12-17, 2019.

[16] H. Byeon, Developing a random forest classifier for predicting the depression and managing the health of caregivers supporting patients with Alzheimer's disease. Technology and Health Care, vol. 27, no. 5, pp. 531-544, 2019.

[17] Ministry of Gender Equality & Family, A study on the national survey of multicultural families. Ministry of Gender Equality & Family, Seoul, 2012.

[18] V. K. Chauhan, K. Dahiya, and A. Sharma, Problem formulations and solvers in linear SVM: a review. Artificial Intelligence Review, vol. 52, no. 2, pp. 803-855, 2019.

[19] D. Li, L. Xu, E. D. Goodman, Y. Xu, and Y. Wu, Integrating a statistical background-foreground extraction algorithm and SVM classifier for pedestrian detection and tracking. Integrated Computer-Aided Engineering, vol. 20, no. 3, pp. 201-216, 2013.

[20] A. Sarica, A. Cerasa, and A. Quattrone, Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: a systematic review. Frontiers in Aging Neuroscience, vol. 9, pp. 329, 2017.

[21] I. A. Ibrahim, T. Khatib, A. Mohamed, and W. Elmenreich, Modeling of the output current of a photovoltaic grid-connected system using random forests technique. Energy Exploration & Exploitation, vol. 36, no. 1, pp. 132-148, 2018.

[22] E. K. Sahin, I. Colkesen, and T. Kavzoglu, A comparative assessment of canonical correlation forest, random forest, rotation forest and logistic regression methods for landslide susceptibility mapping. Geocarto International, vol. 35, no. 4, pp. 341-363, 2020.

[23] J. Rodriguez, L. Kuncheva, and C. Alonso, Rotation forest: a new classifier ensemble method. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 10, pp. 1619-1630, 2006.

[24] E. K. Sahin, I. Colkesen, and T. Kavzoglu, A comparative assessment of canonical correlation forest, random forest, rotation forest and logistic regression methods for landslide susceptibility mapping. Geocarto International, vol. 35, no. 4, 341-363, 2020.

[25] H. Li, Z. Zhang, and Z. Liu, Application of artificial neural networks for catalysis: a review. Catalysts, vol. 7, no. 10, pp. 306. 2017.

[26] R. Adhikari, A neural network based linear ensemble framework for time series forecasting. Neurocomputing, vol. 157, pp. 231-242, 2015.

[27] Y. Ren, L. Zhang, and P. N. Suganthan, Ensemble classification and regression-recent developments, applications and future directions. IEEE Computational Intelligence Magazine, vol. 11, no. 1, pp. 41-53, 2016.

[28] F. Divina, A. Gilson, F. Gomez-Vela, M. Garcia Torres, and J. F. Torres, Stacking ensemble learning for short-term electricity consumption forecasting. Energies, vol. 11, no. 4, pp. 949, 2018.

[29] R. Saini, and S. Ghosh, Ensemble classifiers in remote sensing, a review 2017 International Conference on Computing, Communication and Automation (ICCCA), IEEE, pp. 1148-1152, 2017.

[30] Y. Xiao, J. Wu, Z. Lin, and X. Zhao, A deep learning-based multi-model ensemble method for cancer prediction. Computer Methods and Programs in Biomedicine, vol. 153, pp. 1-9, 2018.

[31] L. Feng, Y. Li, Y. Wang, and Q. Du, Estimating hourly and continuous ground-level PM2.5 concentrations using an ensemble learning algorithm: the ST-stacking model. Atmospheric Environment, vol. 223, pp. 117242, 2020.

[32] J. Chen, J. Yin, L. Zang, T. Zhang, and M. Zhao, Stacking machine learning model for estimating hourly PM2. 5 in China based on Himawari 8 aerosol optical depth data. Science of The Total Environment, vol. 697, pp. 134021, 2019.

[33] S. Lee, and H. Kim, A new ensemble machine learning technique with multiple stacking. Society for e-Business Studies, vol. 25, no. 3, pp. 1-13, 2020.