G-Tech : Jurnal Teknologi Terapan
Volume 7, No. 4, Oktober 2023, hal. 1616-1624
E-ISSN: 2623-064x | P-ISSN: 2580-8737

Modelling The Nexus between Parenting Style and Anti Social Behavior using Ensemble Learning Approach

Angga Aditya Permana[1]✉, Muhammad Fahrury Romdendine[2]

[1, 2] Department Informatics, Faculty of Engineering and Informatics, Universitas Multimedia Nusantara, Indonesia

Informasi Artikel

Riwayat Artikel

Diserahkan : 28-09-2023
Direvisi : 02-10-2023
Diterima : 06-10-2023

ABSTRAK

Masyarakat kontemporer menghadapi masalah perilaku anti-sosial pada anak dan remaja, yang salah satunya dipengaruhi oleh tipe pengasuhan oleh orang tua. Penelitian ini menggunakan teknologi *machine learning*, khususnya *ensemble learning*, untuk memodelkan hubungan antara tipe pengasuhan dengan perilaku anti-sosial. Data penelitian berasal dari studi sebelumnya yang mencakup parameter tipe pengasuhan dan perilaku anti-sosial. Data tersebut dilakukan praproses dan rekayasa fitur, lalu digunakan dalam pemodelan dengan menggunakan metode Random Forest (RF) dan Adaptive Boost (AdaBoost). Pemodelan dilakukan dalam dua tahap: *vanilla modelling*, dan *hyperparemeter tuning*. Hasil model yang telah di-*tuning* menunjukkan bahwa RF lebih baik (akurasi=91%) daripada AdaBoost (akurasi=72%). Dapat disimpulkan bahwa RF sebagai *bagging ensemble learning* mampu memodelkan hubungan antara tipe pengasuhan dan perilaku anti-sosial dengan baik. Studi berikutnya harus disarankan untuk menghimpun lebih banyak data latih dan mengembangkan sistem deteksi dini yang dapat digunakan oleh psikolog anak di lapangan.

Kata Kunci:

Ensemble learning, perilaku anti-sosial, sistem deteksi dini, tipe pengasuhan

ABSTRACT

Contemporary society is grappling with issues of anti-social behavior in children and adolescents, one of which is influenced by parenting styles. This research employs machine learning technology, particularly ensemble learning, to model the relationship between parenting styles and anti-social behavior. The research data is derived from previous studies encompassing parenting style parameters and anti-social behavior. This data is preprocessed and feature-engineered, then used in modelling through the Random Forest (RF) and Adaptive Boost (AdaBoost) methods. Modelling is conducted in two phases: vanilla modelling and hyperparameter tuning. The results of the tuned models indicate that RF performs better (accuracy=91%) than AdaBoost (accuracy=72%). In conclusion, RF, as a bagging ensemble learning technique, effectively models the relationship between parenting styles and anti-social behavior. Future studies are recommended to gather more training data and develop an early detection system for use by child psychologists in the field.

Keywords **:**

Anti-social behavior, early detection system, ensemble learning, parenting style

Corresponding Author **:**
Angga Aditya Permana
Department Informatics, Faculty of Engineering and Informatics, Universitas Multimedia Nusantara
Jalan Scientia Boulevard Gading, Curug Sangereng, Serpong, Kabupaten Tangerang, Banten 15810
Email: angga.permana@umn.ac.id

INTRODUCTION

Contemporary society is currently surrounded by issues related to anti-social behavior, particularly in children and adolescents, necessitating a comprehensive understanding of this issue and the search for preventive solutions across various sectors (Shan et al., 2023). The development of anti-social behavior in children and adolescents is closely linked to their ability to self-control, a capacity significantly influenced by the parenting style they receive from their parents (Pérez-Fuentes et al., 2019). Therefore, it is imperative to further explore the potential relationships between parental parenting styles and the development of anti-social behavior in children.

Several prior studies have addressed the issue of the development of anti-social behavior in children from various perspectives, including those attributed to the influence of social media (Gruzd et al., 2023), the child's relationship with the family (Lott, 2009), and parenting style (Tehrani & Yamini, 2020a). However, to date, no study has employed machine learning to model any of these aspects. The potential outcomes of utilizing machine learning in modelling these aspects could be highly impactful in early detection efforts concerning the emergence of anti-social behavior in children (Rusli et al., 2020). In our research, we specifically focus on modelling the relationship between parenting style and anti-social behavior due to the limited availability of field data pertaining to this aspect alone.

One preventive measure that can be taken to shield children from developing anti-social behavior as early as possible is to predict the types of anti-social behaviors a child might develop based on their parenting style. In contemporary times, machine learning and artificial intelligence technologies have revolutionized various fields due to their increasingly robust predictive capabilities, particularly in the realm of ensemble learning. Ensemble learning offers the advantage of producing reliable, accurate, and transparent models even when the amount of data available is limited since it use decision tree as base learner (Tobing et al., 2019). This is achieved through ensemble learning mechanisms that prevent overfitting while enhancing generalization and reducing variance (Lee et al., 2019).

Machine learning, in general, has been utilized in several previous studies to model the psychological development of children and adolescents. For instance, Chan et al., (2023) employed feed-forward neural networks to predict conduct disorder in adolescents based on various factors. Uchida et al., (2022) utilized ensemble learning as a model for early detection of bipolar development in children. Furthermore, in a review study, the potential use of AI technology, including machine learning, for modelling pediatric behavioral development was discussed (Aylward et al., 2023). To the best of the researcher's knowledge, there has not been a study that leverages machine learning to model the non-linear relationship between parenting style and the social behaviors that children and adolescents may exhibit.

This research aims to model the non-linear relationship between parenting style and several predictor variables such as age and cultural factors (represented by a country's individualism score) concerning various types of anti-social behaviors exhibited by children and adolescents. The modelling will be conducted using ensemble learning, comprising two techniques: bagging and boosting. The bagging technique will be represented by the Random Forest (RF) algorithm, while the boosting technique will be represented by the Adaptive Boosting (AdaBoost) algorithm. Both techniques will be compared, and the best-performing one will be selected. This study holds the potential to serve as a foundation for further research on modelling the psychological development of children and adolescents based on social factors. Additionally, the models generated in this research can also be valuable as the core model for an intelligent system aimed at early detection of anti-social behaviors in children and adolescents.

MATERIALS AND METHODS

The Dataset

The dataset utilized in this research was obtained from data published by Tehrani & Yamini, (2020b). This dataset originated from a study by Tehrani & Yamini, (2020a) aimed at elucidating the relationship between effective child-rearing practices, low self-control, and anti-social behavior within the framework of general strain theory. The researchers obtained this dataset through a meta-analysis study, which involved compiling samples from various papers that measured populations.

The raw data consists of 462 rows and 34 columns. Each column represents features that store data values. Explanations for each feature can be found in the original source publication by Tehrani and Yamini (2020b). Based on the explanations provided in the original publication, only four predictor features and one target feature relevant to the future modelling were selected. The four predictor features are "Mean_Age," which denotes the average age of the sample. "Female_Percentage," representing the percentage of female gender in the sample. "Individualism_Score," indicating the individualism score of the sample based on the country. "Kind_Parenting," specifying the type of parenting practiced by the parents in the sample. Meanwhile, "Kind_ASB" is the target feature, signifying the type of antisocial behavior exhibited by the sample.

Data Preprocessing and Feature Engineering

Before using the acquired data for machine learning, it must be preprocessed and engineered. This involves handling missing values, ensuring numerical values are standardized, and converting categorical features into numerical ones using one-hot encoding. These steps ensure that the data is in a suitable format for effective machine learning modelling.

Handling missing values aims to prevent bias and inaccuracies in machine learning models (Gond et al., 2021). In this study, due to the relatively low number of missing values, imputation was chosen as the method of handling missing data. Imputation involved replacing missing values with the median for numerical data and the mode for categorical data. Additionally, standardization was applied to the numerical predictor features, "Mean_Age" and "Female_Percentage," due to the different units of measurement in these features. This ensures that the data is prepared appropriately for accurate machine learning modelling.

The encoding of categorical predictor features into numerical values is performed because machine learning models typically find it easier to learn patterns from numerical data (Staartjes & Kernbach, 2022). Given that ensemble learning is a part of machine learning, feature engineering is crucial to achieve good model performance. One-hot encoding is chosen as the method for encoding categorical features because it introduces no bias to the model. This is because one-hot encoding encodes each categorical feature into binary values, ensuring there are no order implications (Zhuang et al., 2021).

Ensemble Learning Modelling and Hyperparameter Tuning

In this research, two ensemble learning algorithms, RF (Breiman, 2001) and AdaBoost (Hastie et al., 2009), were employed. These two algorithms were chosen because they represent two different ensemble learning techniques. RF represents the bagging technique, while AdaBoost represents the boosting technique. This choice allows for an assessment of which ensemble learning technique is more effective in modelling non-linear relationships between various predictor features and the target feature.

Bagging, also known as Bootstrap Aggregating, is an ensemble learning strategy that includes training many models on distinct subsets of the training data and then aggregating their predictions to generate a final prediction. Bagging is employed in the case of Random Forest to generate numerous decision trees, each trained on a random subset of the training data and a random subset of the features. The ultimate prediction is then created by averaging all of the trees' predictions. This strategy aids in reducing overfitting and improving model accuracy and stability.

Random Forest captures a larger range of patterns in the data and reduces the importance of any given characteristic by employing a random subset of the features for each tree.

Boosting is a popular ensemble-learning technique that combines multiple weak learners to create a strong learner. AdaBoost, short for Adaptive Boosting, is a widely used boosting algorithm that iteratively trains weak learners on different subsets of data and assigns higher weights to misclassified samples in each iteration. The final model is a weighted combination of weak learners, in which the weights are determined by their performance on training data. AdaBoost is particularly effective in classification problems because it can improve the accuracy of the model by reducing bias and variance. One of the main advantages of AdaBoost is its ability to handle high-dimensional and noisy data, making it a popular choice in many real-world applications. However, AdaBoost can be sensitive to outliers and overfitting, which can be mitigated by tuning the hyperparameters or using other boosting algorithms, such as Gradient Boosting or XGBoost. Overall, AdaBoost is a powerful and versatile ensemble learning technique that can significantly improve the performance of machine-learning models.

The modelling process, as previously mentioned, consists of two stages: vanilla modelling and hyperparameter tuning. Vanilla modelling involves building models without any modifications or, in other words, using default hyperparameters. This stage is aimed at obtaining the base performance of each algorithm. Base performance is crucial during hyperparameter tuning to determine whether there is an improvement in performance after modifying the hyperparameters of each algorithm. Performance evaluation, both during vanilla modelling and after hyperparameter tuning, is conducted using common metrics for multi-class classification cases, including accuracy, macro F1-Score, micro F1-Score, weighted F1-Score, and one vs. one (OVO) ROC-AUC score.

In multiclass classification, evaluating the performance of a model is crucial to determine its effectiveness in predicting the correct class labels. Accuracy, macro f1-score, micro f1-score, and OVO ROC-AUC score are some of the commonly used metrics to evaluate the performance of a multiclass classification model. Accuracy is the most straightforward metric that measures the percentage of correctly classified instances out of the total number of instances (Nurul Rismayanti & Aulia Putri Utami, 2023). However, accuracy can be misleading when the dataset is imbalanced, and the number of instances in each class is not equal.

Macro f1-score is a metric that calculates the harmonic mean of precision and recall for each class and then takes the average across all classes. It is useful when the dataset is imbalanced, and the number of instances in each class is not equal (Tran et al., 2022). Macro f1-score gives equal weight to each class, regardless of the number of instances in each class. Micro f1-score, on the other hand, calculates the harmonic mean of precision and recall across all classes, taking into account the total number of true positives, false positives, and false negatives. Micro f1-score is useful when the dataset is balanced, and the number of instances in each class is equal.

OVO ROC-AUC score is a metric that measures the ability of a model to distinguish between all pairs of classes. It is calculated by training a binary classifier for each pair of classes and then computing the area under the receiver operating characteristic (ROC) curve for each binary classifier. OVO ROC-AUC score is useful when the dataset has more than two classes and the classes are not mutually exclusive (Nguyen et al., 2023).

Hyperparameter tuning is performed using a random search technique, implemented using RandomizedSearchCV available in the Scikit-Learn package in the Python programming language. RandomizedSearchCV searches for hyperparameters based on a configured search space provided as a dictionary. RandomizedSearchCV also performs cross-validation when searching for the best hyperparameters for the model. In this study, a 10-fold cross-validation is employed. The hyperparameter search space configurations for both algorithms are presented in Table 1.

Table 1. Hyperparameter Configuration used in Model Tuning

Algorithm	Hyperparameter Configuration
RF	{ 'n_estimators': [100, 200, 500], 'criterion': ['gini', 'entropy', 'log_loss'], 'max_depth': [None, 5, 10, 20], 'min_samples_split': [2, 5, 10], 'min_samples_leaf': [1, 2, 4] }
AdaBoost	{ 'n_estimators': [50, 100, 200], 'learning_rate': [0.01, 0.1, 1], 'algorithm': ['SAMME', 'SAMME.R'] }

RESULTS AND DISCUSSION

Model Performances

Vanilla modelling for RF and AdaBoost yielded different results, with RF generally outperforming. The confusion matrices for the vanilla RF and AdaBoost models are presented in Figure 1. Accuracy metrics from the evaluation of the vanilla models are presented in Table 2. Although not achieving very high performance, an accuracy of 72% is still a respectable result. Considering the limited training data and the fact that these are still vanilla models, these results indicate good performance.

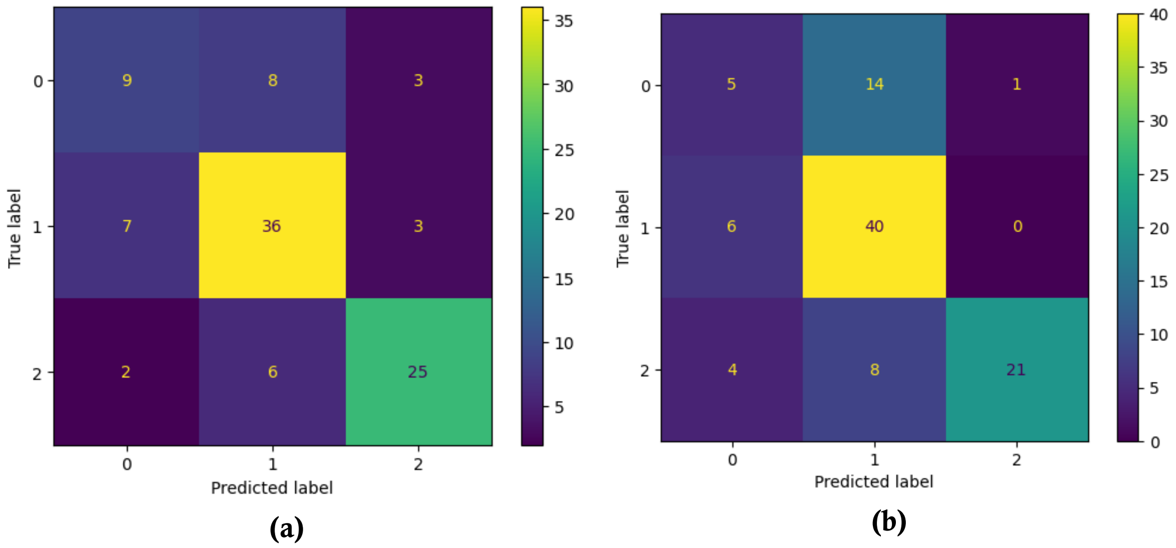


Figure 1. Confusion Matrix of Vanilla RF (a); and Vanilla AdaBoost (b).

The label of "0" indicates a group of anti-social behavior consists of vandalism, theft, and assault, group fight, shot or stabbed someone, and pulled a knife or a gun on someone, physical assault, shoplifting, carry a hidden weapon, and attack someone. The label of "1" indicates a group of anti-social behavior consists of alcohol use, school misconduct, sell drugs, write bad checks, gang membership, nonviolent crime, substances use, childhood antisociality, risky lifestyles, running away home, and risk-taking behaviors. The lable of "2" indicates general deviance behavior.

The performance metrics obtained during the vanilla modelling process serve as the base performance for RF and AdaBoost. This base performance is used as a reference for comparing the results of hyperparameter tuning for both algorithms. The confusion matrices for tuned RF and tuned AdaBoost are presented in Figure 2. The performance of the models after hyperparameter tuning is documented in Table 3.

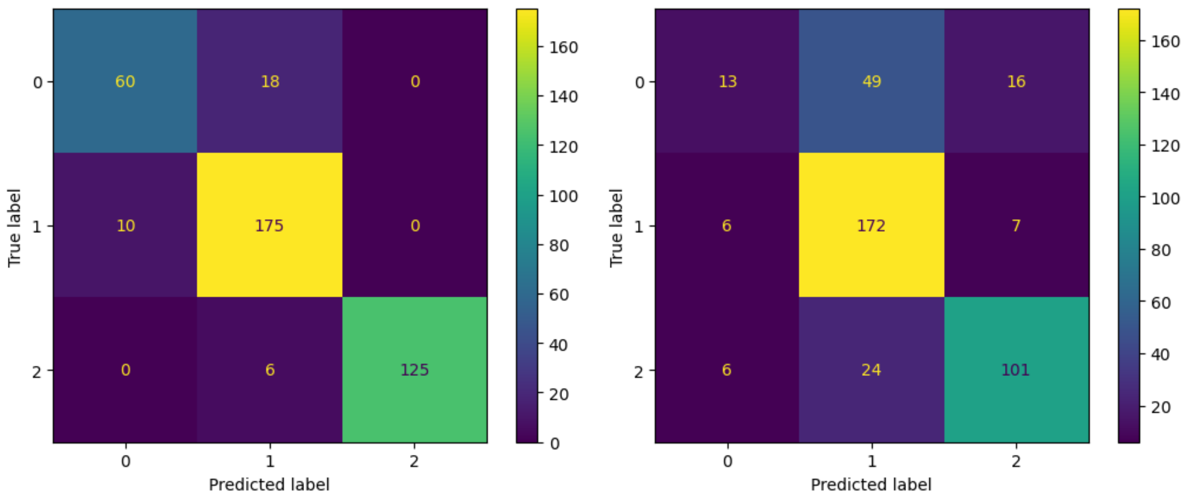
Tabel 2. Evaluation Metrics of Vanilla Models from each Algorithm

Metric	RF	AdaBoost
Accuracy	0.71	0.67
Macro F1-Score	0.67	0.60
Micro F1-Score	0.71	0.67

Metric	RF	AdaBoost
Weighted F1-Score	0.70	0.66
ROC-AUC Score	0.87	0.82

Tabel 3. Evaluation Metrics of Tuned Models from each Algorithm

Metric	RF	AdaBoost
Accuracy	0.91	0.72
Macro F1-Score	0.90	0.61
Micro F1-Score	0.91	0.73
Weighted F1-Score	0.91	0.69
ROC-AUC Score	0.99	0.83



(a) **(b)**

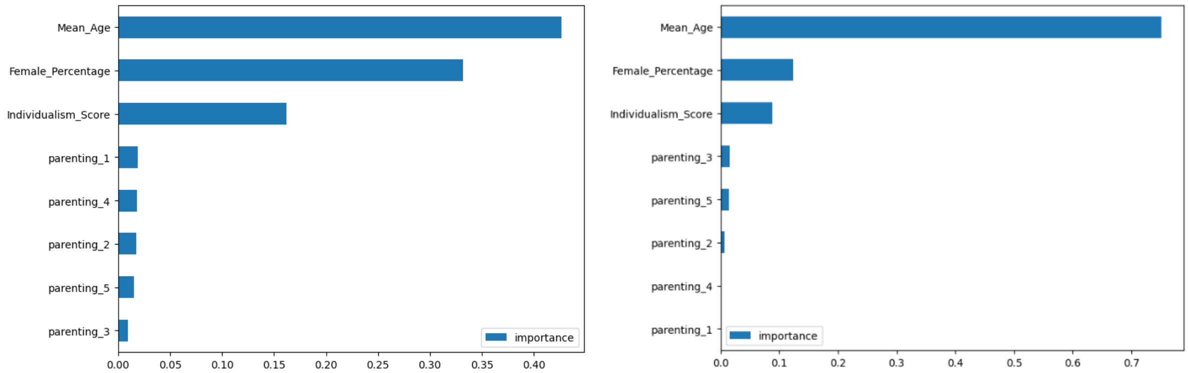
Figure 2. Confusion Matrix of Tuned RF (a); and Tuned AdaBoost (b).

The performance of both algorithms significantly improved after hyperparameter tuning using random search. There was a substantial improvement in the RF algorithm, with an increase in accuracy by 20%, macro F1-Score by 23%, micro F1-Score by 20%, weighted F1-Score by 21%, and ROC-AUC score by 12%. Meanwhile, the performance of AdaBoost improved with an accuracy increase of 5%, macro F1-Score by 1%, micro F1-Score by 6%, weighted F1-Score by 3%, and ROC-AUC score by 1%.

The significant difference in performance between RF, which is a bagging ensemble learning method, and AdaBoost indicates that the bagging mechanism is more effective in modelling the non-linear relationship between parenting styles and antisocial behavior. Bagging is an ensemble learning technique that leverages multiple weak classifiers to make predictions through a voting mechanism, where the majority vote determines the predicted class (Ngo et al., 2022). The superiority of bagging (RF) over boosting (AdaBoost) in our experiments can be attributed to the fact that bagging (RF) has the capability to learn from data containing a lot of noise and is less prone to overfitting (due to the relatively smaller dataset used) (Yildirim et al., 2018).

Model Analysis

The results of the feature importance analysis for both models indicate that "Mean_Age," which describes the age of the sample, is the most influential predictor feature in determining the possibility of antisocial behavior occurring in the sample. Both models also rank "Female_Percentage," which represents the percentage of females in the sample, and "Individualism_Score," describing the individualism score of the country where the sample is from, as the second and third most important features. This suggests that the percentage of female gender and the level of individualism significantly influence the type of antisocial behavior exhibited by the sample. However, there are differences in the significance scores of the parenting types that have been encoded into values from 1 to 5. The feature importance analysis plot is presented in Figure 3.



(a) (b)

Figure 3. Feature Importance Scores from RF model (a); and from AdaBoost (b)

CONCLUSIONS

Based on the obtained results, several conclusions can be drawn. The bagging technique in ensemble learning, represented by the RF algorithm, demonstrated its superiority in modelling the non-linear relationship between parenting style and various other predictor features with antisocial behavior. Additionally, the hyperparameter tuning process using random search significantly improved the model's performance, especially for RF. However, the differing order of significance given by both algorithms to each parenting style does not conclusively prove a significant influence of parenting style on the type of antisocial behavior observed in the sample. Nonetheless, this does not dismiss the potential impact of parenting style entirely. This variation in significance might be due to the limited data available from the source, compounded by the encoding of each parenting style into binary form, resulting in sparse matrices.

There are still limitations in this study that could be addressed in future research. The limited amount of training data used in this study may have affected the robustness of the model analysis compared to real-world observations. Future research can aim to gather a larger dataset to enhance the credibility and applicability of the models. Nevertheless, the findings from this research can serve as a foundation for developing an intelligent system that could be utilized by parenting practitioners as a tool for observing the potential development of antisocial behavior in a child based on the parenting style they receive from their parents.

ACKNOWLEDGMENT

The authors would like to thank Universitas Multimedia Nusantara for the support of this research work.

REFERENCE

Aylward, B. S., Abbas, H., Taraman, S., Salomon, C., Gal-Szabo, D., Kraft, C., Ehwerhemuepha, L., Chang, A., & Wall, D. P. (2023). An Introduction to Artificial Intelligence in Developmental and Behavioral Pediatrics. *Journal of Developmental and Behavioral Pediatrics : JDBP*, *44*(2), e126–e134. https://doi.org/10.1097/DBP.0000000000001149

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chan, L., Simmons, C., Tillem, S., Conley, M., Brazil, I. A., & Baskin-Sommers, A. (2023). Classifying Conduct Disorder Using a Biopsychosocial Model and Machine Learning Method. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *8*(6), 599–608.

https://doi.org/10.1016/j.bpsc.2022.02.004

Gond, V. K., Dubey, A., & Rasool, A. (2021). A Survey of Machine Learning-Based Approaches for Missing Value Imputation. *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*, 1–8. https://doi.org/10.1109/ICIRCA51532.2021.9544957

Gruzd, A., Soares, F. B., & Mai, P. (2023). Trust and Safety on Social Media: Understanding the Impact of Anti-Social Behavior and Misinformation on Content Moderation and Platform Governance. *Social Media + Society*, *9*(3). https://doi.org/10.1177/20563051231196878

Hastie, T., Rosset, S., Zhu, J., & Zou, H. (2009). Multi-class AdaBoost. *Statistics and Its Interface*, *2*(3), 349–360. https://doi.org/10.4310/SII.2009.v2.n3.a8

Lee, J., Kim, J., & Ko, W. (2019). Day-Ahead Electric Load Forecasting for the Residential Building with a Small-Size Dataset Based on a Self-Organizing Map and a Stacking Ensemble Learning Method. *Applied Sciences*, *9*(6), 1231. https://doi.org/10.3390/app9061231

Lott, R. E. (2009). *The unique influence of mothers' and fathers' negative parental practices on adolescent anti-social behavior: Mediating effects of adolescent psychological resources and deviant peer relationships* [Iowa State University, Digital Repository]. https://doi.org/10.31274/etd-180810-549

Ngo, G., Beard, R., & Chandra, R. (2022). Evolutionary bagging for ensemble learning. *Neurocomputing*, *510*, 1–14. https://doi.org/10.1016/j.neucom.2022.08.055

Nguyen, N. D., Tan, W., Du, L., Buntine, W., Beare, Ri., & Chen, C. (2023). AUC Maximization for Low-Resource Named Entity Recognition. *Proceedings of the AAAI Conference on Artificial Intelligence*, *37*(11), 13389–13399. https://doi.org/10.1609/aaai.v37i11.26571

Nurul Rismayanti, & Aulia Putri Utami. (2023). Improving Multi-Class Classification on 5-Celebrity-Faces Dataset using Ensemble Classification Methods. *Indonesian Journal of Data and Science*, *4*(2), 124–133. https://doi.org/10.56705/ijodas.v4i2.78

Pérez-Fuentes, M. del C., Molero Jurado, M. del M., Gázquez Linares, J. J., Oropesa Ruiz, N. F., Simón Márquez, M. del M., & Saracostti, M. (2019). Parenting Practices, Life Satisfaction, and the Role of Self-Esteem in Adolescents. *International Journal of Environmental Research and Public Health*, *16*(20), 4045. https://doi.org/10.3390/ijerph16204045

Rusli, A., Suryadibrata, A., Nusantara, S. B., & Young, J. C. (2020). A Comparison of Traditional Machine Learning Approaches for Supervised Feedback Classification in Bahasa Indonesia. *IJNMT (International Journal of New Media Technology)*, *7*(1), 28–32. https://doi.org/10.31937/ijnmt.v1i1.1485

Shan, Q., Yu, X., & Tian, Y. (2023). Adolescent social isolation shifts the balance of decision-making strategy from goal-directed action to habitual response in adulthood via suppressing the excitatory neurotransmission onto the direct pathway of the dorsomedial striatum. *Cerebral Cortex*, *33*(5), 1595–1609. https://doi.org/10.1093/cercor/bhac158

Staartjes, V. E., & Kernbach, J. M. (2022). Foundations of Machine Learning-Based Clinical Prediction Modeling: Part III-Model Evaluation and Other Points of Significance. *Acta Neurochirurgica. Supplement*, *134*, 23–31. https://doi.org/10.1007/978-3-030-85292-4_4

Tehrani, H. D., & Yamini, S. (2020a). Parenting practices, self-control and anti-social behaviors: Meta-analytic structural equation modeling. *Journal of Criminal Justice*, *68*, 101687. https://doi.org/10.1016/j.jcrimjus.2020.101687

Tehrani, H. D., & Yamini, S. (2020b). Dataset of parenting practices, self-control and anti-social behaviors: Meta-analytic structural equation modeling. *Data in Brief*, *32*, 106114. https://doi.org/10.1016/j.dib.2020.106114

Tobing, F. A. T., Dzulhaq, M. I., Overbeek, M. V., & Setiawan, E. A. (2019). Design of Decision

Support System with C4.5 Algorithm Methods in Determining Results Vaname Shrimp Cultivation. *2019 5th International Conference on New Media Studies (CONMEDIA)*, 195–200. https://doi.org/10.1109/CONMEDIA46929.2019.8981848

Tran, T., Le, U., & Shi, Y. (2022). An effective up-sampling approach for breast cancer prediction with imbalanced data: A machine learning model-based comparative analysis. *PLOS ONE*, *17*(5), e0269135. https://doi.org/10.1371/journal.pone.0269135

Uchida, M., Bukhari, Q., DiSalvo, M., Green, A., Serra, G., Hutt Vater, C., Ghosh, S. S., Faraone, S. V, Gabrieli, J. D. E., & Biederman, J. (2022). Can machine learning identify childhood characteristics that predict future development of bipolar disorder a decade later? *Journal of Psychiatric Research*, *156*, 261–267. https://doi.org/10.1016/j.jpsychires.2022.09.051

Yildirim, P., Birant, K. U., Radevski, V., Kut, A., & Birant, D. (2018). Comparative analysis of ensemble learning methods for signal classification. *2018 26th Signal Processing and Communications Applications Conference (SIU)*, 1–4. https://doi.org/10.1109/SIU.2018.8404601

Zhuang, J., Liu, D., Lin, M., Qiu, W., Liu, J., & Chen, S. (2021). PseUdeep: RNA Pseudouridine Site Identification with Deep Learning Algorithm. *Frontiers in Genetics*, *12*, 773882. https://doi.org/10.3389/fgene.2021.773882