A Stacked Ensemble Framework through Synergistic Modeling for Adolescent Delinquency Classification

Abhinash Jenasamanta *, Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi, India. Corresponding Author Email: samantajena99@outlook.com

Subrajeet Mohapatra, Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi, India. Email: subrajeets@gmail.com

Abstract: This study emphasizes how ensemble approaches can improve the precision and resilience of predictive models in assessing social problems like adolescent delinquency. To classify adolescent delinquency, this study looks into how effective a stacked ensemble framework is. The performance of the stacked ensemble model is compared to individual classifiers using data on different risk factors linked to delinquent behavior. The evaluation metrics that are used are AUC, F-score, and the area under the receiver operating characteristic curve (AUC). The stacked ensemble framework constantly outperforms individual classifiers, as evidenced by the high AUC scores it achieves, which range from 0.97 to 1.00 across all classes of delinquency severity. Moreover, in terms of classification accuracy and F-score, the proposed ensemble model outperforms individual classifiers, demonstrating greater discriminative power. Additionally, the stacked ensemble exhibits improved generalization and robustness, proving its effectiveness in locating underlying patterns in the data and lowering the chance of overfitting. The classification of adolescent delinquency is significantly impacted by the ensemble framework's superior performance, which implies that ensemble techniques like stacking can improve predictability and accuracy in identifying adolescents who are at risk at an early stage. The results demonstrate the potential of the suggested stacked ensemble approach to enhance predictive performance and provide reliable classification across a range of delinquency severity levels, thereby validating its efficacy in classifying juvenile delinquency.

Keywords: Adolescent delinquency, early stage identification, stacked ensemble framework, classification.

INTRODUCTION

Using automated methods, data mining [1] is a technique for obtaining useful information from large data sets. By utilizing machine learning and advanced analytics, it is very easy to develop methods and procedures to categorize data for use in practical applications. Machine learning is now a commonly used technology in modern research and for applications in the medical profession. Classification, regression, and clustering are the most widely used learning techniques for pattern identification [2]. The study of computer algorithms that get better over time by using their expertise and data is known as machine learning (ML). To generate predictions or judgments without being explicitly trained to do so, machine learning algorithms build a model from "training data." Deviance behavior, often known as conduct disorder, is a psychiatric illness that affects adolescents and is characterized by a pattern of persistent rule-breaking and antisocial behavior [3]. Deviant behavior can be identified by the following signs viz. violence or grave threats to harm persons or animals and Consistently lying to escape consequences or to get material goods or privileges. Juveniles are those who are under the age of eighteen. Minors who intentionally or unintentionally commit crimes are considered juvenile delinquents.

Adolescents who engage in juvenile delinquency frequently experience psychological disorders. Some circumstances drive adolescents' behavior to go out of control, which leads to crime. The juvenile justice system has spent a lot of time and energy attempting to understand the root causes of misbehavior [4]. Many frameworks explain how variables and results are related. Juvenile delinquency does not have a single cause, according to researchers, but the presence of several risk factors together significantly increases the likelihood that a child would commit a crime. In the current study, we concentrate on using behavioral, socio-demographic, and machine-learning principles for the early prediction of juvenile delinquency. The above risk variables are gathered using a standard psychometric questionnaire International Self-Report Delinquency Study (ISRD-3) and a socio-demographic assessment to evaluate the child's current psychological condition [5]. The machine learning module (stacked ensemble framework) will use this set of weighted feature scores to divide adolescent behavior into three categories: low, medium, and high. A machine learning model will be created using this compiled set of variables. Following that, the adolescent will be categorized as having a low, medium, or severe risk of engaging in delinquent behavior. High-profile incidents of young people committing

 ISSN: 0017-2294

violent crimes have increased in the modern world. Determining the risk and protective factors is therefore critical for the early detection of juvenile offenders. After closely examining the research directions in the body of existing literature, it has been determined that there is sufficient room to develop an automated system for the early screening and quantitative assessment of delinquent behavior in adolescents. The overall steps of the procedure followed are shown in Fig 1.

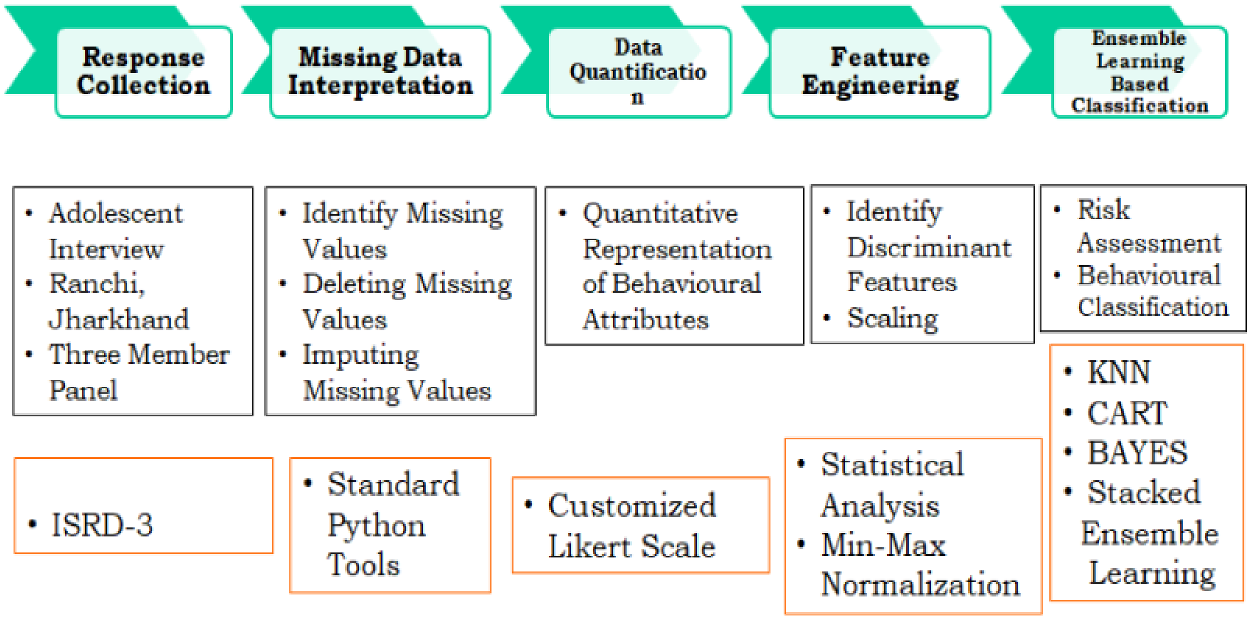

Fig 1. Overall procedures for applying the suggested framework for automated assessment of delinquency in adolescents.

LITERATURE SURVEY

When recent years have witnessed an increase in violent crime incidents involving children. As a result, identifying factors for the early detection of child delinquents is critical [6]. Each of the factors responsible for deviant behavior also needs to be analyzed separately using quantitative data analysis [7]. Following is the analysis of literature explaining the various factors responsible for deviant behavior in juveniles. A study conducted using multivariate analysis of variance (MANOVA) showed that the deficiency of proper nutrition leads to multiple behavioral issues in children viz. aggression and conduct disorder [8]. Food insecurity, social interaction, and neurocognitive behavior were assessed in both male and female participants. Children pointing towards malnutrition resulted in improper social behavior (p-value less than 0.0001) when compared to children with sufficient nutritional values. Regression analysis was conducted to study the relationship between stressful changes in life patterns and their impact on five different types of crime and delinquency [9]. Regression analysis showed that, for both males and females, stress and life pattern change are highly related to the prediction of delinquent activities. Tobit regression analysis along with descriptive statistics was implemented to find out whether moral factors have a relationship with offending in the Netherlands [10].

Their findings revealed that moral feelings had a strong and direct influence on criminal behavior for juveniles. A longitudinal study of children below twenty-one examined the connection between early physical violence and extreme forms of deviant activities occurring in the later period by using ANCOVA (Analysis of Covariance) [11]. According to their findings, people who faced physical forms of abuse during childhood were more likely to be arrested as juvenile offenders in the later period of life. A Structural equation model (SEM) to analyze the effect of economic strain on delinquency in adolescents was developed [12]. They analyzed the data of 244 families consisting of 122 younger brothers and an equal number of younger sisters. The results proved that sibling aggression has a very strong and harmful impact on delinquency in adolescents who come from economically strained families.

Factor analysis and logistic regression using a three-level random intercept model were implemented with control for clustering to study the effects of student perceptions of the school environment on delinquency [13]. They discovered that the understanding of meaningful schoolwork can mitigate the negative effects of wrongful conditions at home for boys, such as family members' alcohol consumption.

Correlation analysis and multi-level ordinal regression modeling were carried out to verify if the neighborhood factors influence delinquency and drug use while keeping the impact of individual behavioral effects as the control on the youth of age sixteen [14]. They found out that the neighborhood attributes of the youth's residence have a less significant influence on their delinquency and drug use habits. Latent growth curve modeling (LGM) along with the weighted least squares means-variance was implemented to analyze the relationship between adolescent delinquency and alcohol consumption with unlawful acts of adults, early alcohol addiction, and risky sex [15]. The fitness of the Model was calculated by using the Tucker–Lewis Index (TLI) and the Root Mean Square Error (RMSE). Late childhood delinquency had an indirect relationship with youth violence, alcohol disorders, and risky sex, according to their report, but it has a continuous direct impact on crimes. An extensive study on Queensland children using logistic regression was conducted which showed that aggression at the age of five is the most important predictor of aggression at the age of 14. They also concluded that aggression at the early age of preschool is a very important factor in predicting future delinquency and suspension from school [16]. A simulation model that works in tandem with a Geographic Information System (GIS) to assess various social policies aimed at reducing crime in urban areas was implemented [17]. They tested the model with wildly disparate socioeconomic indices. Because of the successful interaction of the GIS, simulation model, and database, the established model may support complex decision-making.

METHODOLOGY

An attempt has been made in the current work to assess adolescent behavioral risk factors for delinquency using a stacking framework. In this section, we outline the specifics of the suggested system's design as well as the methodical procedure used. Fig. 1 depicts the overall procedure of classification. After identifying the research, completing a review of the literature, and formulating goals and hypotheses for the study, a methodical approach was created. The study's focus was on the Indian state of Jharkhand because of the high percentage of poverty there and the recent arrests of many young people for crimes.

DATA COLLECTION:

125 adolescents enrolled in classes VII through XII of the chosen schools were interviewed to gather the primary data. Secondary information on each person was gathered by interviewing the relevant teachers and included information about his prior social behavior in the school. ISRD-3 India Version, questionnaire was used for the interview, along with the appropriate explanation based on regional conditions. Participants in this study ranged in age from 12 to 17 and were from the outer Ranchi area of Jharkhand, India. The principals, who oversee each school, were contacted first, and then the students in the classroom, to obtain the necessary consent to take part in the survey.

Data were gathered in Ranchi, Jharkhand, from 13 schools (public and private) by a three-person team comprising the authors and a behavioral psychologist. Each school's head of faculty is permitted to conduct a survey there. Later, it was asked of the students in the classroom if they would agree to take part in the survey. Lastly, the survey team had in-person interviews with the students.

DATA CLEANING:

The collected dataset has no missing data.

FEATURE GENERATION:

Computer-aided delinquency diagnosis requires accurate characterization of adolescent behavior traits. Using guidelines provided by a two-member clinical psychologist panel, the survey set was quantified and transformed into 47 features. For a given factor/feature measurement, every data field has been transformed into a corresponding numerical value. The two-member panel recommended using standard psychological scaling techniques viz. Likert Scale to measure environmental, parental, and individual factors [18]. In addition to the previously mentioned elements, victimization and gang affiliations are additional features that are examined, quantified, and graphically depicted for computer-based analysis. The final generated features are provided in the table below along with their respective short forms:

Table 1: Table of Attributes

ACTUAL ATTRIBUTES	RESPECTIVE SHORT FORM
urban/rural, Gender, Age, ethnicity, no. of parents, mother tongue, belief, importance of religion,	atr1, atr2, atr3, atr4, atr5, atr6, atr7, atr8
Category, income source, poverty, parent attachment, parental supervision, Family Disturbance, school bond, School environment,	atr9, atr10, atr11, atr12, atr13, atr14, atr15, atr16.
Teacher attachment, Unexplained Absentism, Academic Performance, future plans, Victimization, going out, subjective wellbeing, foreign friends, Leisure activities, Peer Attachment, Delinquent Peers	atr17, atr18, atr19, atr20, atr21, atr22, atr23, atr24, atr25, atr26, atr27
Behavior towards delinquency, perception of shame, Aggresion, Traumetic Injury, Neighbour association, neigbouring environment, Offending, Animal Cruelty, Arrest Record, Alcohol use, Drug abuse, Substance Abuse Integrity	atr28, atr29, atr30, atr31, atr32, atr33, atr34, atr35, atr36, atr37, atr38, atr39
Norm strength, Procedural justice, Gang, age group of peers, Acceptance of Illegal Activities, Delinquent group Peers, response integrity, researcher observation	atr40, atr41, atr42, atr43, atr44, atr45, atr46, atr47

In future sections , figures and calculations we have used the short forms of generated attributes for ease of use.

DATA VISUALIZATION:

The individual data features are visualized in figures below:

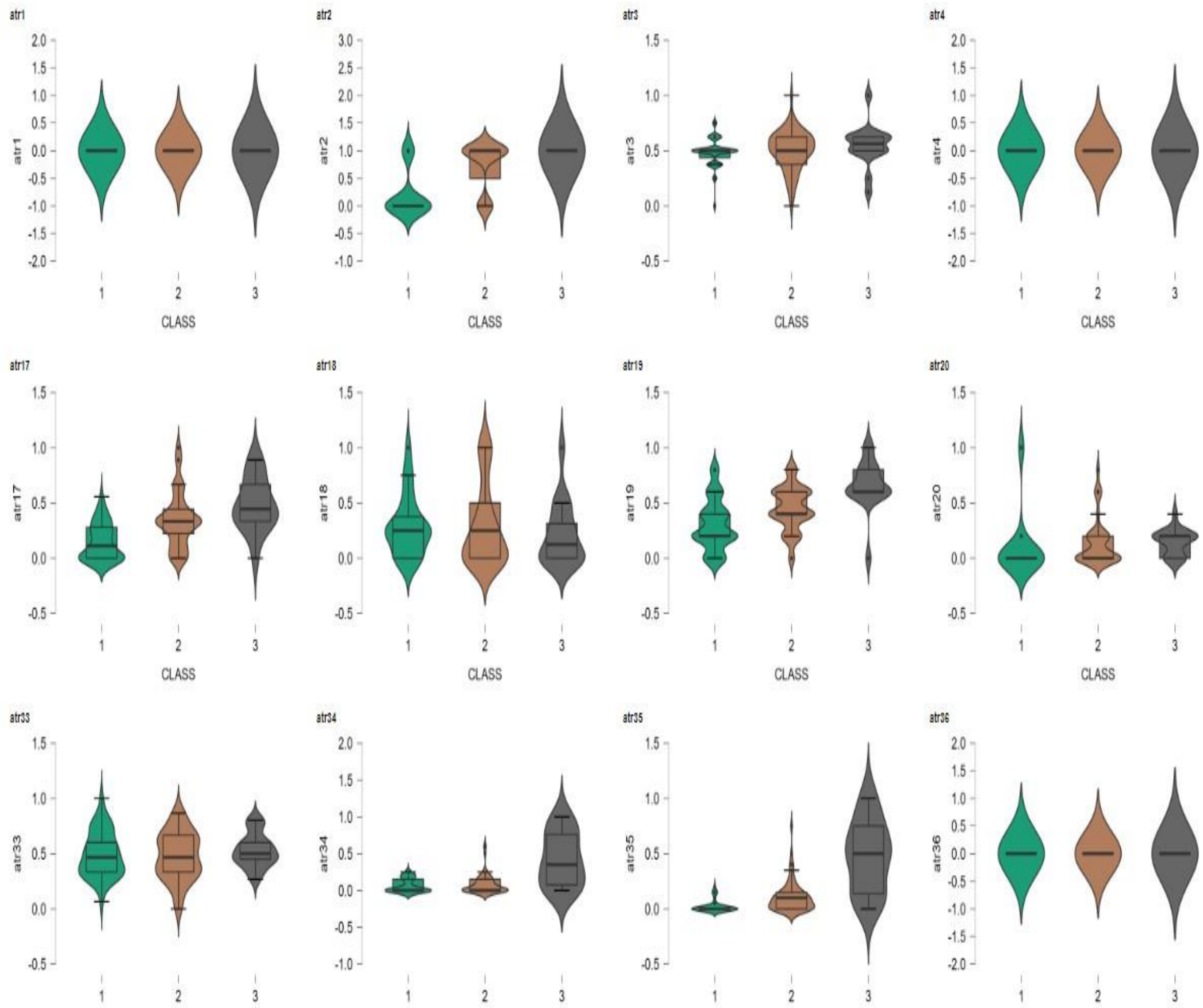

Fig 2: Visualizing individual data features

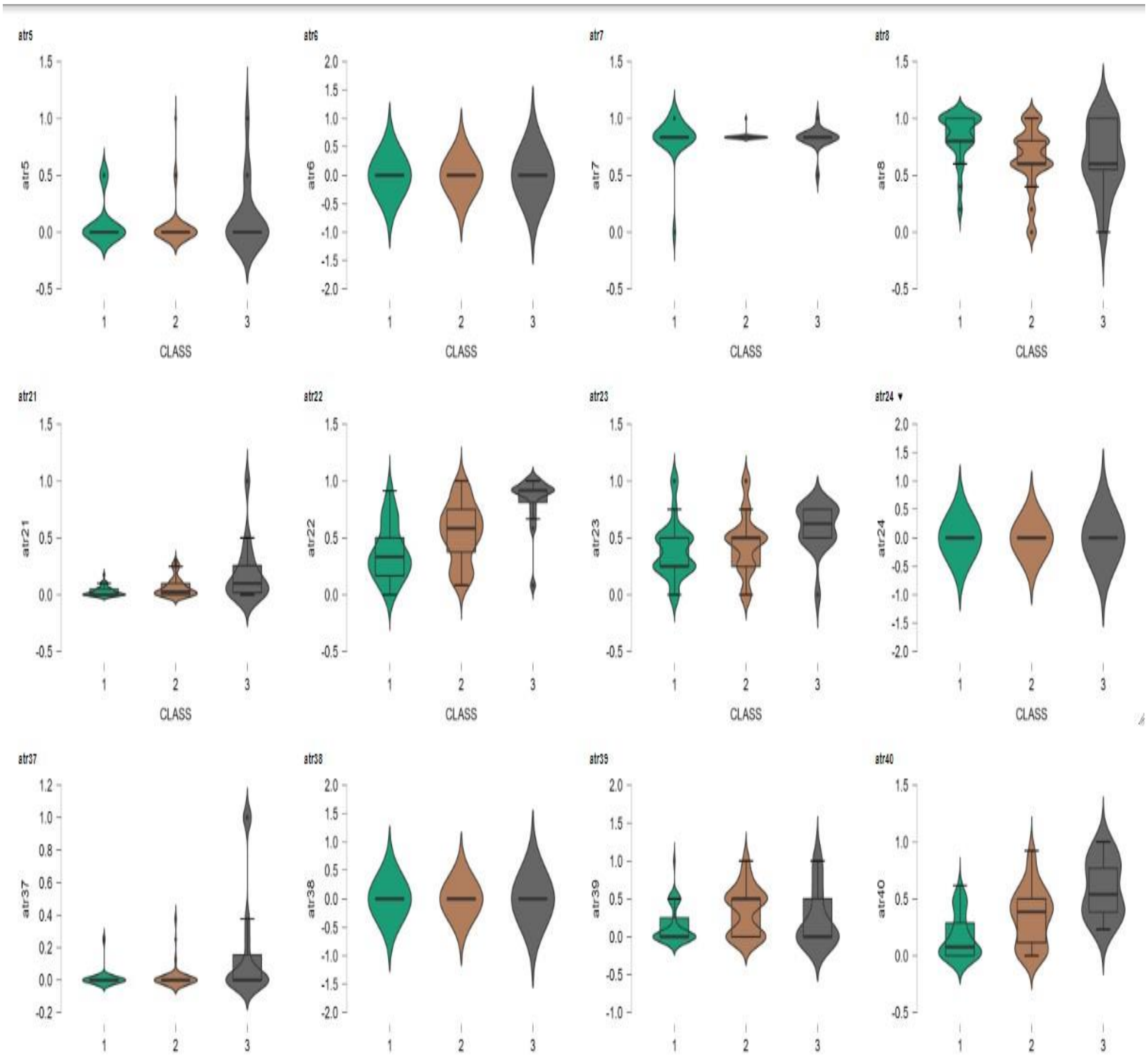

Fig 3: Visualizing individual data features

GORTERIA
2024 64(5)
ISSN: 0017-2294

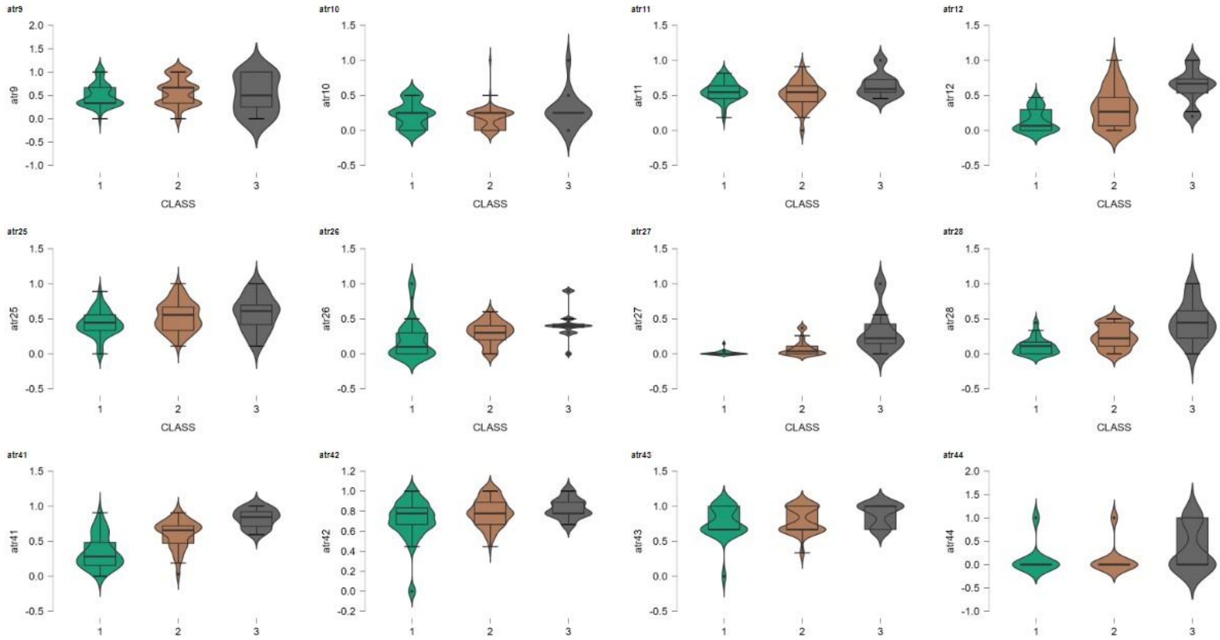

Fig 4: Visualizing individual data features

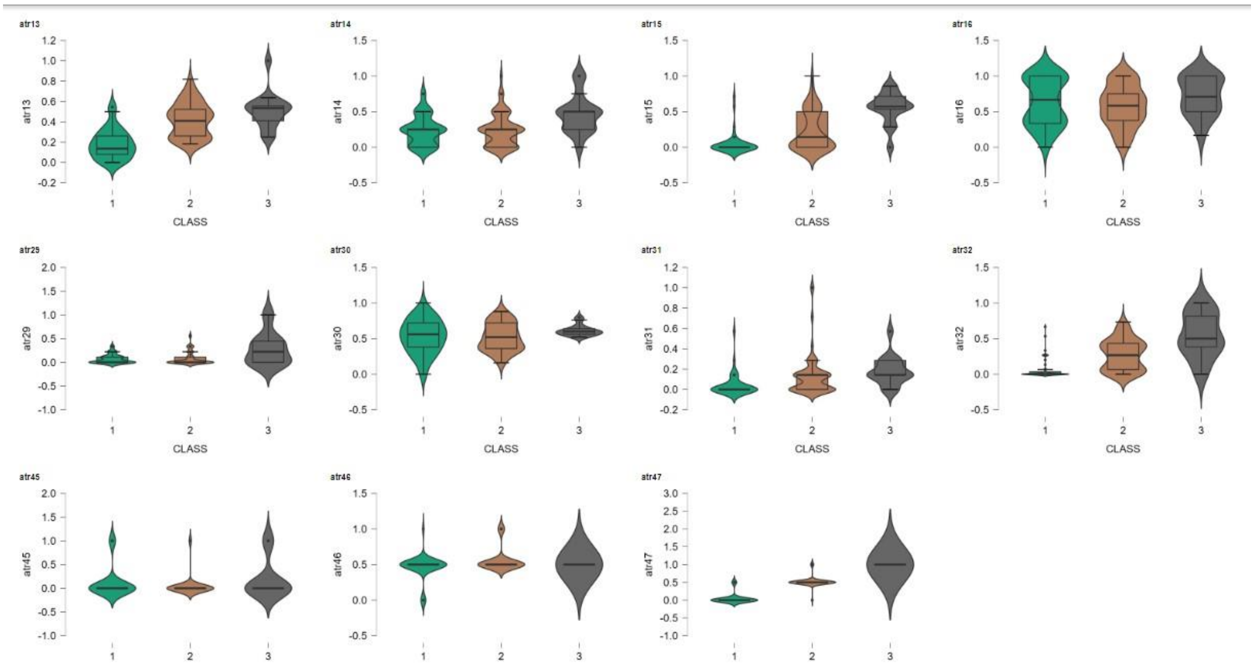

Fig 5: Visualizing individual data features

ISSN: 0017-2294

The heat map showing the various attributes are shown below:

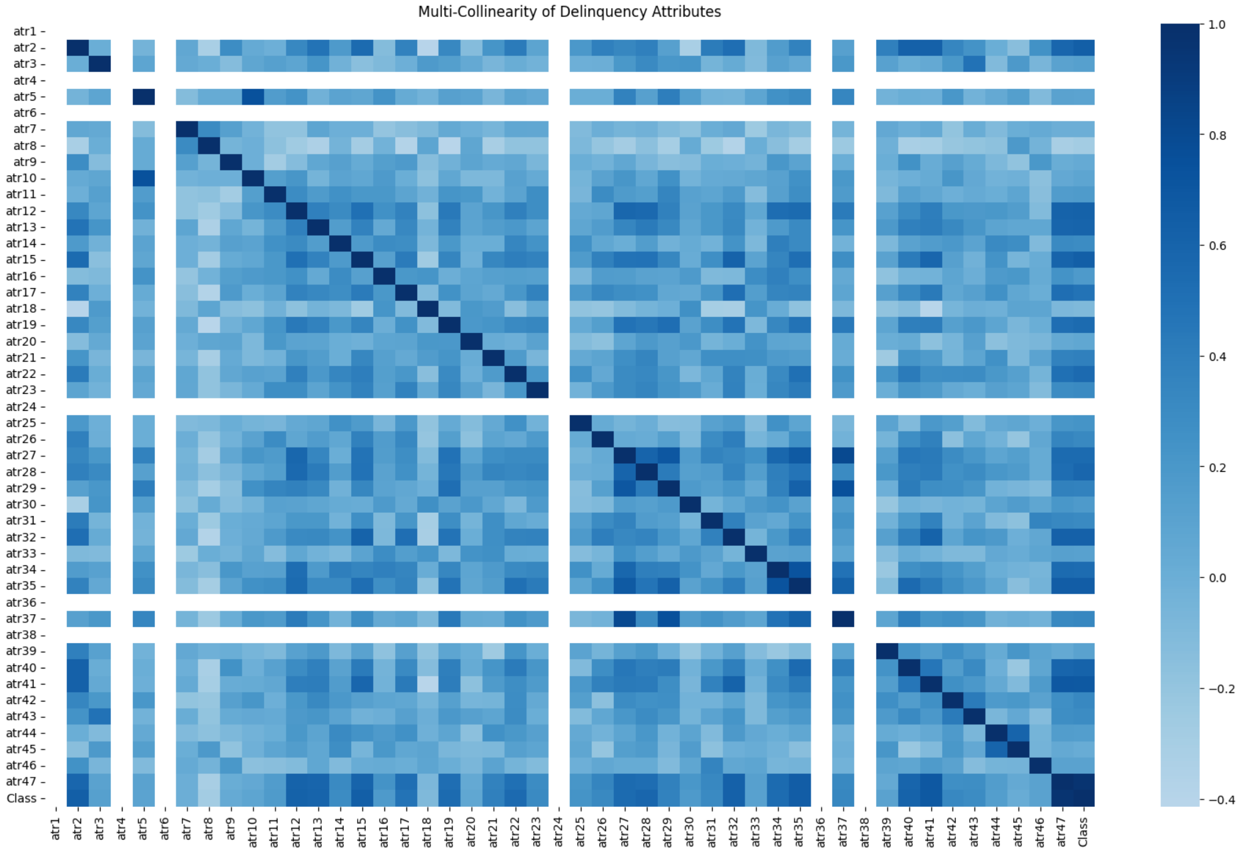

Fig 6: Visualizing correlation between features using Heat Map

FEATURE SCALING USING MIN MAX NORMALIZATION:

Min-Max normalization has been utilized to scale all the feature values within a range of 0 to 1 for standardized analysis [19].

FEATURE IMPORTANCE USING "SELECT K BEST" TECHNIQUE

The best features from a given dataset are extracted using the Scikit-learn Select K Best class. The features are selected by the Select K Best method by utilizing the k highest score [20]. The 'score_func' parameter can be changed to apply the method to both regression and classification data. Selecting the best features is an important step in the process of getting ready to train a large dataset. It speeds up the training process and makes it easier to remove less important data.

ISSN: 0017-2294

If we want to find out the best 10 Features in our dataset we adjust the parameters as:

Select_features = Select K Best (score_func = f_classif, k=10) # Select top 10 features

The result is :

Selected Feature Indices: [1 11 12 21 27 31 33 34 40 46]
Accuracy: 0.9615384615384616

The above result means that the best 10 important features are 1,11,12,21,27,31,33,34,40 and 46.

So by changing the value of 'k' we can get as many best features out of the total features.

RELIABILITY ANALYSIS USING GUTTMAN'S LAMBDA

To evaluate a set of variables' ability to discriminate between various groups or factors within a dataset, one statistical metric that is used is Guttman's Lambda (λ) [21]. In discriminant analysis, factor analysis, and related multivariate statistical techniques, it is especially pertinent. The amount of variance in the dependent variable(s) (such as factor scores or group membership) that can be explained by the independent variables (such as predictor variables and observed variables) is known as Guttman's Lambda. The range of Guttman's Lambda is 0 to 1, where: perfect discrimination is represented by a value of 1, which means that all group or factor variance in the dependent variable(s) is fully explained by the variables. poor discrimination is indicated by a value near zero, which shows that the variables are ineffective in differentiating between groups or factors. Higher values indicate stronger discrimination; values between 0 and 1 represent varying degrees of discrimination.

It is calculated in Python code as follows:

```
# Calculate covariance matrix

cov_matrix = np.cov(numeric_data, rowvar=False)

# Calculate the total variance
```

ISSN: 0017-2294

```
total_variance = np.sum(np.var(numeric_data, axis=0))

# Calculate Guttman's Lambda

lambda_value = np.trace(cov_matrix) / total_variance
```

RESULT: Guttman's Lambda: 1.008

So we got the Guttman’s Lambda value as 1.008.

When Guttman's Lambda values are close to 1, it means that the variables are good at differentiating between groups or factors; when it is much less than 1, it means that the discrimination is weaker [22]. Efficient differentiation among groups or factors enables us to more fully comprehend and describe the variations or contrasts found in the data. More accurate prediction and grouping of observations into distinct categories are made possible by discriminative variables. Finding variables that successfully separate groups or factors aids in focusing on and prioritizing the most important characteristics for additional research or analysis. Stratified analysis—in which distinct subgroups within the dataset are examined independently to identify patterns or relationships unique to each subgroup—is made possible by discriminative variables.

CLASSIFICATION

KNN

For supervised learning, K-Nearest Neighbors (KNN) is a popular supervised machine learning technique [23]. It is a non-parametric, instance-based approach, therefore it does not have any fundamental assumptions on how the data are distributed. Instead, it creates predictions based on how closely related the data points are. KNN determines the K closest neighbors to a new, unknown data point in the training dataset based on a similarity measure (often Euclidean distance, though other distance metrics can be used as well). We can select K's value as a hyperparameter. In KNN, the selection of the distance measure is essential. Euclidean distance is the most widely used, although depending on the nature of the problem and the data, alternative distance metrics, such as Manhattan distance, Minkowski distance, or custom distance functions, may be utilized.

CART

The decision tree algorithm CART (Classification and Regression Trees) is used in machine learning for both classification and regression applications [24]. It is a potent and understandable technique that creates a structure like a tree to make predictions based on input information. CART's primary principle is to repeatedly divide the dataset into subsets according to the values of various attributes until a stopping requirement is satisfied. Starting with the complete dataset, the CART algorithm takes into account all features that are available and the matching target variable. To divide the dataset into two or more subsets, CART chooses the most advantageous characteristic. It assesses the quality of each feature using a metric like mean squared error (for regression) or Gini impurity (for classification). After the tree has been constructed, it can be pruned to eliminate branches that don't significantly improve the model's ability to predict the future. Overfitting can be avoided with pruning.

NAÏVE BAYES:

A probabilistic machine learning technique called Naive Bayes is used for classification and, occasionally, regression [25]. It relies on the "naive" assumption that the features are conditionally independent, which simplifies the computation of probabilities. It is based on the Bayes theorem. Naive Bayes typically performs unexpectedly well in a variety of real-world applications, such as the classification of spam emails and document classification, despite its simplicity and naive assumption. The Bayes theorem, which connects the likelihood of an event occurring given certain data to its prior likelihood, serves as the cornerstone of Naive Bayes. It aids in determining the likelihood of a specific class label given the features in a classification context.

ENSEMBLE LEARNING USING STACKING:

The proposed stacking ensemble model for the classification of adolescent delinquency is shown below in Fig 7. The base classifiers are KNN, CART (Decision Tree), and Naïve Bayes. By implementing the stacking technique we get the final prediction using the below architecture.

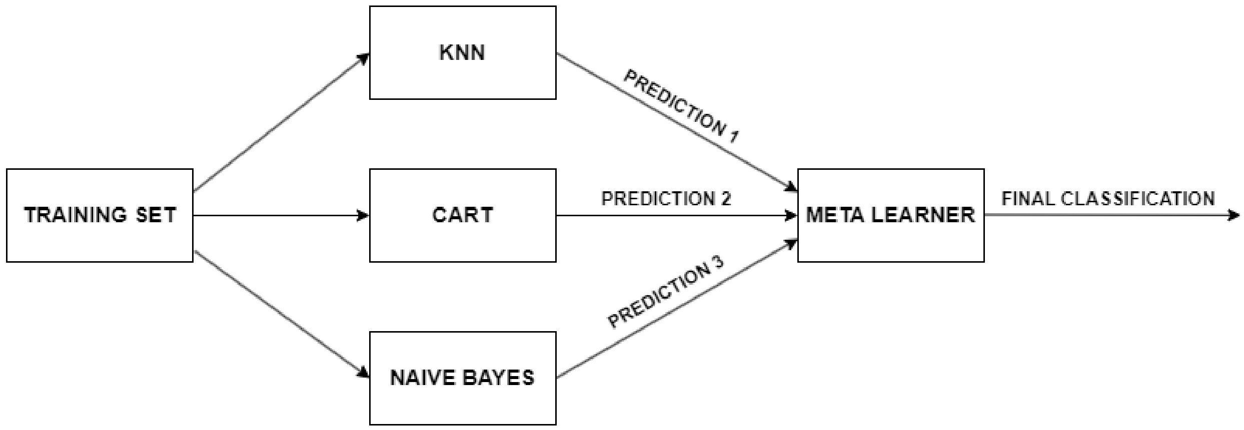

Fig 7.Proposed Stacked Ensemble Framework

A stacking ensemble is a machine-learning method that combines multiple models to enhance prediction performance [26]. It is also referred to as stacking or stacked generalization. Essentially, the goal is to train a meta-model that discovers the optimal way to integrate the base models' predictions.

The steps of the Stacking algorithm are explained below:

Step 1- Base Models: To begin, train a collection of varied base models with various algorithms or variations on the same algorithm.

Step 2-Verification Information: Divide your training set into two subsets: the subset used to train the base models and the subset used to generate the meta-features. Cross-validation is a useful tool for this.

Step 3-Base Model Forecasts: Utilizing the first portion of the training data to train each base model, forecast the second portion of the data.

Step 4-Meta-Model: The base models' predictions are now your new dataset, with each row representing an instance from the initial training set and each column representing a distinct base model. Apply the actual target values from the training data as labels to this new dataset and train a meta-model.

Step 5-Forecasting: After the meta-model has been trained, it can be used to forecast new data by first producing predictions with the base models and then feeding those predictions into the meta-model.

EVALUATION PARAMETERS

Cross-Validation:

When there is little data available, K-fold cross-validation is used to get a better evaluation of the model's performance [27]. K number of equal-sized subsets of the data are created.
We use one of the training subsets as the test set and build K models, omitting one each time. This is known as a "Leave-One-Out" if K equals the sample size.

Classification Accuracy: One of the metrics used to assess a classification model's performance is classification accuracy [28].

Classification Accuracy = Number of Correct Predictions / Total Predictions

F1 SCORE:

A popular metric for classification tasks that strikes a balance between recall and precision is the F1 score [29].

F1 = 2* [(Precision * Recall) / (Precision + Recall)]

The F1 score has a maximum value of 1 (perfect recall and precision) and a minimum value of 0. When we want to strike a balance between recall and precision, this metric comes in handy.

RESULTS AND DISCUSSION:

The training and validation of the model are done using repeated stratified K-Fold cross-validation [30]. The model performance results of the individual classifiers are compared with the proposed stacking framework.

Table 2: Comparative performance comparison of various classifiers with the proposed stacking framework.

METHODOLOGY	CLASSIFICATION ACCURACY	F-SCORE
KNN	88.70%	0.88
CART	96 %	0.96
NAÏVE BAYES	92.30%	0.92
PROPOSED STACKED ENSEMBLE MODEL	97.40 %	0.97

The graphical comparison of the performance of proposed framework with individual base classifiers is shown in figure below:

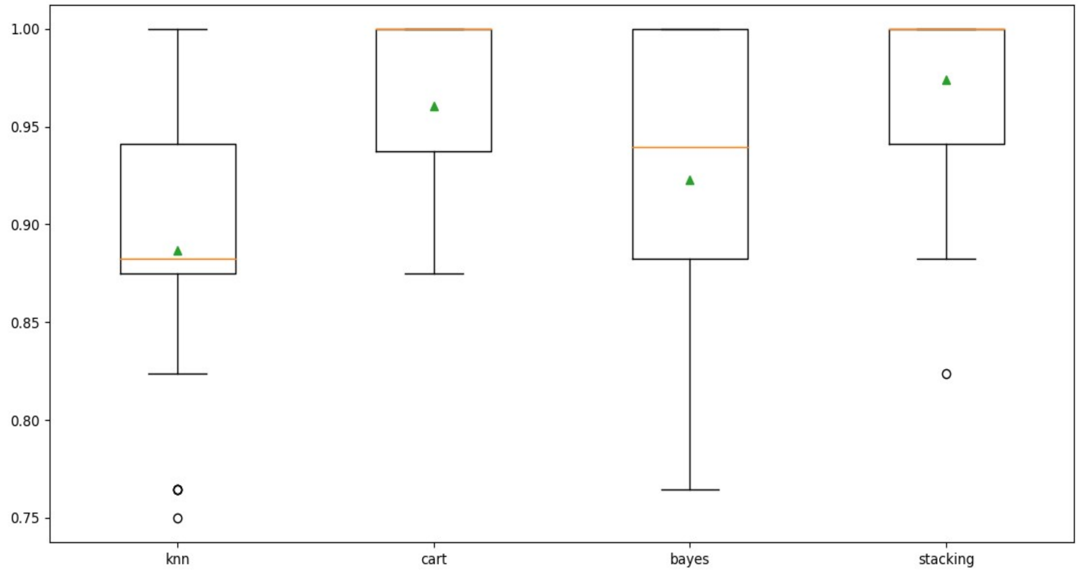

Fig 8. Graphical comparison of the performance of the proposed framework with individual base classifiers.

ROC CURVE AND AREA UNDER THE CURVE (AUC) ANALYSIS:

The AUC-ROC curve is used to measure the performance of the classification problems at various threshold settings [31]. A probability curve is called a ROC, and an AUC is a measure of separability [32]. It shows how well the model can distinguish between different classes. AUC values close to 1 indicate a high-quality model with good separability. AUC value nearing 0 indicates a poor model, which has the least separability [33]. Additionally, the model has no capacity for class separation when the AUC is 0.5.

ROC CURVE FOR KNN:

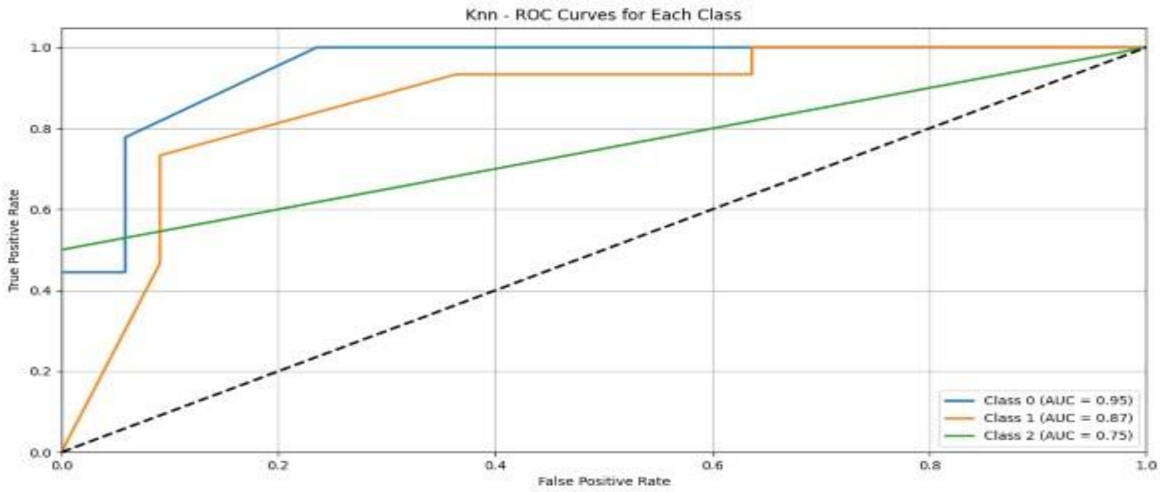

Fig 9. ROC Curve for KNN

CART ROC CURVE:

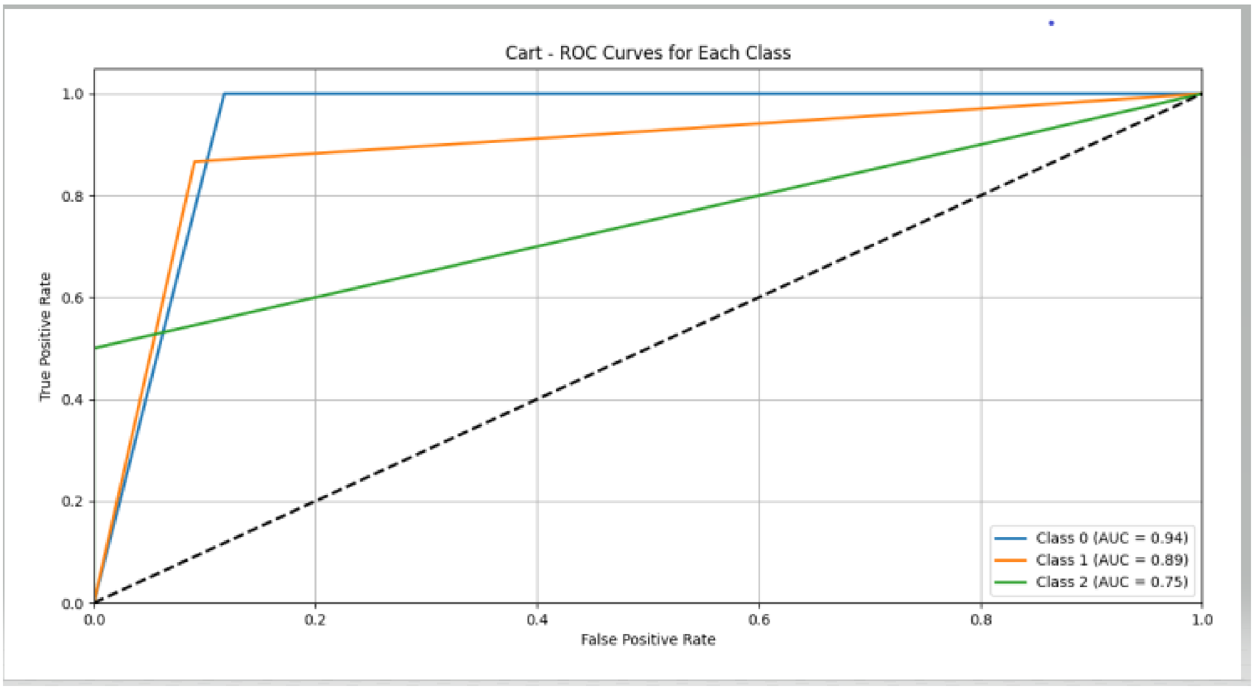

Fig 10. ROC Curve for CART

NAÏVE BAYES ROC:

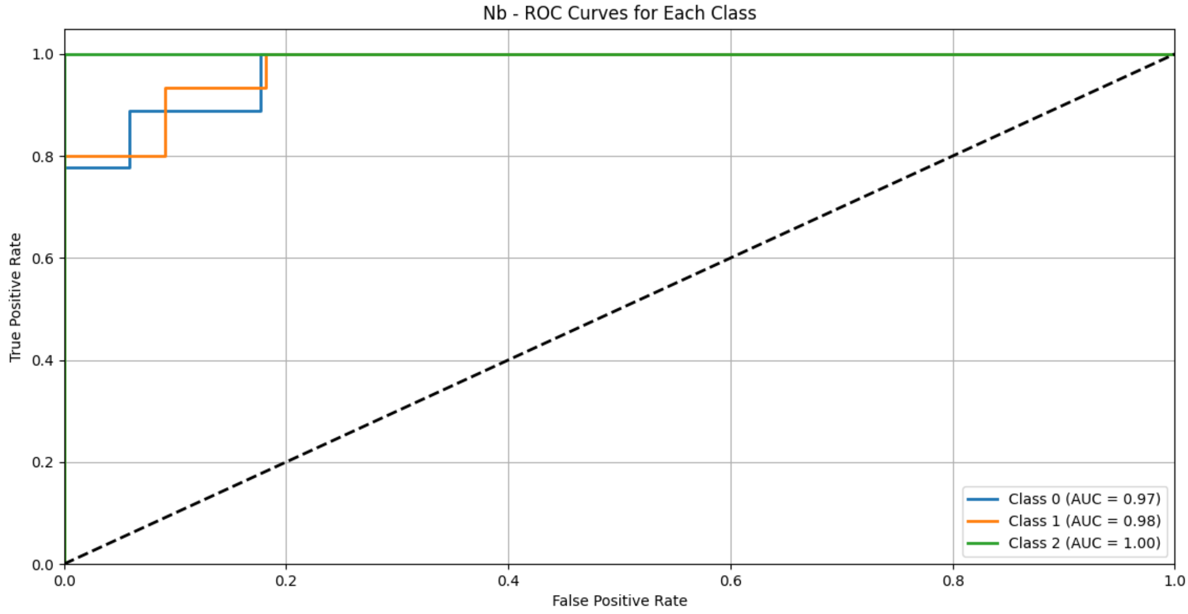

Fig 11. ROC Curve for Naïve Bayes

ROC CURVE FOR PROPOSED STACKED ENSEMBLE FRAMEWORK::

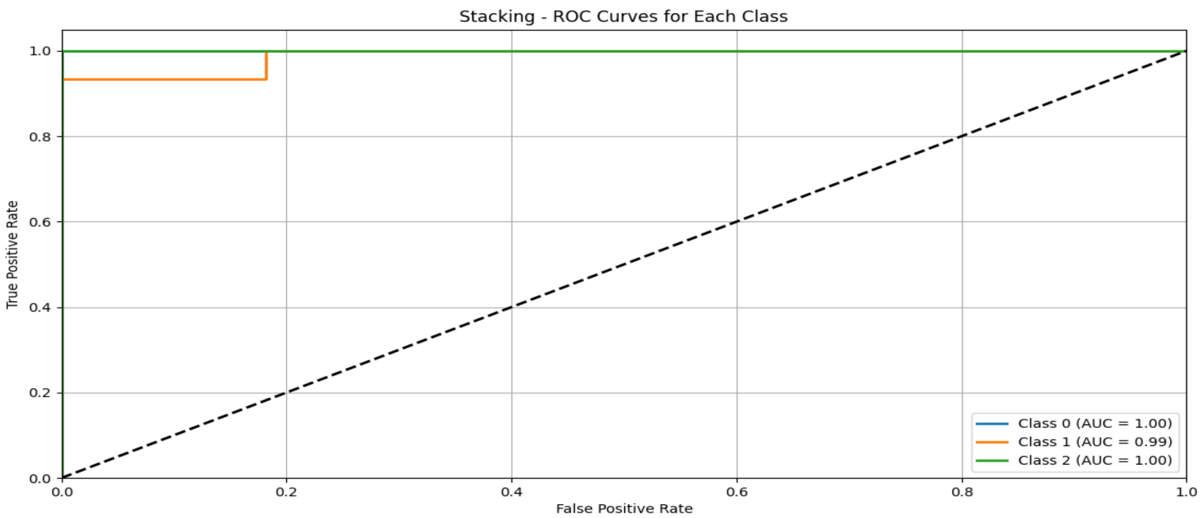

Fig 12. ROC Curve for Proposed Stacked Ensemble Framework

COMPARATIVE ROC CURVE ANALYSIS SHOWING ALL THE IMPLEMENTED CLASSIFIERS AND PROPOSED STACKING FRAMEWORK:

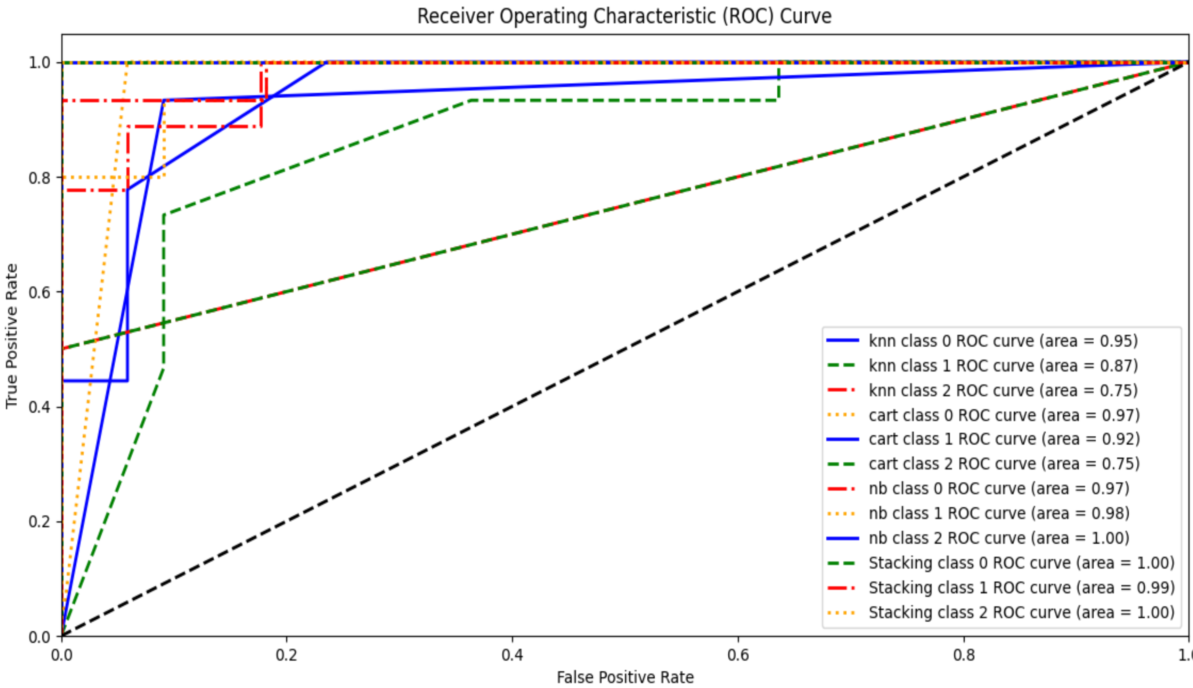

Fig 13. Comparative ROC curve analysis showing all the implemented classifiers and proposed stacking framework:

Table 3: Comparative performance analysis of various classifiers with our proposed model

CLASSIFIERS	(AUC)		
	CLASS 0(LOW)	CLASS 1(MED)	CLASS 2 (HIGH)
KNN	0.95	0.87	0.75
CART	0.97	0.92	0.75
NAÏVE BAYES	0.97	0.98	1.00
PROPOSED STACKED ENSEMBLE FRAMEWORK	1.00	0.99	1.00

The following conclusions can be made using the suggested stacked ensemble framework and the supplied table that displays the area under the ROC curve (AUC) for various classifiers:

Performance of Each Classifier Individually:

KNN: Achieves AUC scores ranging from moderate to high in various classes, with class 0 (Low) receiving the highest score of 0.95.

CART: AUC scores of 0.97 and 0.92 for class 0 and class 1, respectively, show strong discriminative ability.

PERFORMANCE OF PROPOSED MODEL: With values ranging from 0.97 to 1.00, the stacked ensemble framework continuously achieves high AUC scores in all classes. When compared to individual classifiers, the ensemble approach exhibits better discriminative power because it combines predictions from several classifiers.

CONCLUSION

The following succinctly describes the conclusion drawn from comparing individual classifiers with the suggested stacked ensemble model for the multi-class classification of adolescent delinquency :

i.) Model Performance: In terms of F-score and classification accuracy, the suggested stacked ensemble model performs better than individual classifiers. With a higher classification accuracy of 97.40%, it outperforms the individual classifiers in terms of overall predictive performance.

ii.) Ensemble Advantage: Predictive accuracy and F-score have increased when multiple classifiers are stacked in an ensemble framework. The ensemble model can effectively leverage diverse modeling approaches to improve predictive performance by combining the strengths of various base classifiers.

iii.) When compared to individual classifiers, the stacked ensemble's better performance indicates greater robustness and generalization capabilities. The ensemble model is more adept at identifying the underlying patterns in the data and is less prone to overfitting because it aggregates predictions from several models.

iv.) The stacked ensemble model's better performance has significant ramifications for real-world classification task applications. It suggests that using ensemble methods, like stacking, can greatly increase the predictive models' accuracy and dependability, resulting in more sensible decisions made in practical situations.

v.) In terms of AUC, the suggested stacked ensemble framework performs better than individual classifiers, demonstrating its potency in classifying instances across all classes. By stacking different classifiers, one can combine their complementary strengths to achieve better classification performance. The ensemble framework's high AUC scores indicate that it can classify data reliably and accurately at various levels of class distinction.

To sum up, the outcomes validate the stacked ensemble framework's suitability for classification tasks by emphasizing its capacity to improve predictive performance and offer dependable classification across various classes. The findings validate the effectiveness of the suggested stacked ensemble method for the classification of adolescent delinquency, showing that it can outperform individual classifiers and offer improved predictive accuracy and robustness.

Accessibility of data: The corresponding author can provide the dataset used for analysis in this study upon reasonable request.

Acknowledgments: The research was funded by the Cognitive Science Research Initiative (CSRI), Department of Science & Technology (DST), Government of India, for which the authors are grateful. The necessary permissions for the use of the ISRD-3 questionnaire were provided by Ineke Haen Marshall, Professor in the Department of Sociology and School of Criminology and Criminal Justice at Northeastern University in Boston, for which the authors are grateful. The psychologist's panel's assistance with various scaling purposes is acknowledged with gratitude by the authors.

ISSN: 0017-2294

Competing interests

The authors declare no competing interests.

Ethical approval

With adherence to all relevant guidelines, both authors have conducted research on " A Stacked Ensemble Framework through Synergistic Modeling for Adolescent Delinquency Classification”**.** This research is not subject to the institute's ethics approval committee because it does not involve any clinical studies involving humans.

Informed Consent:

The principal of each school granted the authors of this study the necessary authorization to conduct student interviews. The authors were also granted the necessary permissions to conduct the survey, from the respective class teachers. Participants were also given information about the research and their rights to participate or not. Participants gave their informed consent for the survey. Additionally, verbal informed consent was obtained over the phone from each participant's parents after they were informed of the study's goals.

REFERENCES:

[1] Dol, Sunita M., and Pradip M. Jawandhiya. "Classification technique and its combination with clustering and association rule mining in educational data mining—A survey.
" Engineering Applications of Artificial Intelligence 122 (2023): 106071.

[2]Farshidvard, A., Farnaz Hooshmand, and S. A. MirHassani. "A novel two-phase clustering-based under-sampling method for imbalanced classification problems." Expert Systems with Applications 213 (2023): 119003.

[3] Umidjon, Yusupov, and Abduvakhidova Dilyora. "Factors Influencing The Formation of Deviant Behavior In Adolescence." Conferencea (2023): 89-93.

[4] Watson, Jennifer, India Bryce, Tania M. Phillips, Tait Sanders, and Annette Brömdal. "Transgender youth, challenges, responses, and the juvenile justice system: A systematic literature review of an emerging literature.
" Youth Justice 24, no. 1 (2024): 88-112.

[5] Marshall, Ineke Haen, Katharina Neissl, and Anna Markina. "A global view on youth crime and victimization: results from the International Self-Report Delinquency Study (ISRD3).
" Journal of contemporary criminal justice 35, no. 4 (2019): 380-385.

[6] Krohn, Marvin D. and Jodi Lane. “The handbook of juvenile delinquency and juvenile justice.” (2015).

[7] “Social-Psychological Factors of Deviant Behavior of Adolescents.” Bulletin of Science and Practice (2021): n. pag.

[8] Liu, Jianghong and Adrian Raine. “The effect of childhood malnutrition on externalizing behavior.” Current Opinion in Pediatrics 18 (2006): 565–570.

[9] Vaux, Alan C. and Mary Ruggiero. “Stressful life change and delinquent behavior.” American Journal of Community Psychology 11 (1983): 169-183.

[10] Svensson, Robert, Frank M. Weerman, Lieven J. R. Pauwels, Gerben J. N. Bruinsma and Wim Bernasco. “Moral emotions and offending: Do feelings of anticipated shame and guilt mediate the effect of socialization on offending?” European Journal of Criminology 10 (2013): 22 - 39.

[11] Lansford, Jennifer E., Shari Miller-Johnson, Lisa J. Berlin, Kenneth A. Dodge, John E. Bates and Gregory S. Pettit. “Early Physical Abuse and Later Violent Delinquency: A Prospective Longitudinal Study.” Child Maltreatment 12 (2007): 233 - 245.

[12]Low, Sabina, Ryan Thomas Sinclair and Joann Wu Shortt. “The role of economic strain on adolescent delinquency: a microsocial process model.” Journal of family psychology : JFP :

ISSN: 0017-2294

journal of the Division of Family Psychology of the American Psychological Association 26 4 (2012): 576-84 .

[13] Sandahl, Julia. "School climate and delinquency – on the significance of the perceived social and learning climate in school for refraining from offending." Journal of Scandinavian Studies in Criminology and Crime Prevention 17 (2016): 110 - 130.

[14] McVie, Susan and Paul A. Norris. "Neighbourhood Effects on Youth Delinquency and Drug Use." (2006).

[15] Mason, W. Alex, Julia E Hitch, Rick Kosterman, Carolyn A. McCarty, Todd I. Herrenkohl and J. David Hawkins. "Growth in adolescent delinquency and alcohol use in relation to young adult crime, alcohol use disorders, and risky sex: a comparison of youth from low- versus middle-income backgrounds." Journal of child psychology and psychiatry, and allied disciplines 51 12 (2010): 1377-85 .

[16] Bor, William, Jake M. Najman, Michael J. O'callaghan, Gail M. Williams and Kaarin Jane Anstey. "Aggression and the Development of Delinquent Behaviour in Children." Trends and issues in crime and criminal justice (2001): 1.

[17] Quijada, Sergio E., J. Arcas, Cristian Renner and Luis Rabelo. "A spatio temporal simulation model for evaluating delinquency and crime policies." Proceedings of the Winter Simulation Conference, 2005. (2005): 7 pp.-.

[18] "Likert Scale in Social Sciences Research: Problems and Difficulties." FWU Journal of Social Sciences (2022): n. pag.

[19] Henderi, Henderi, Tri Wahyuningsih and Efana Rahwanto. "Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer." (2021).

[20] Kramer, Oliver, and Oliver Kramer. "Scikit-learn." Machine learning for evolution strategies (2016): 45-53.

[21] Zeng, Zihao, Liyi Peng, Shuangjin Liu, Qin Yang, Hongcai Wang, Zhen He, and Yiqiu Hu. "Serotonergic multilocus genetic variation moderates the association between interpersonal relationship and adolescent depressive symptoms." Journal of Affective Disorders 340 (2023): 616-625.

[22] Bianchi, Renzo, Jay Verkuilen, James F. Sowden, and Irvin Sam Schonfeld. "Towards a new approach to job-related distress: A three-sample study of the Occupational Depression Inventory." Stress and Health 39, no. 1 (2023): 137-153.

[23] Huang, Anzhong, Rui Xu, Yu Chen, and Meiwen Guo. "Research on multi-label user classification of social media based on ML-KNN algorithm." Technological Forecasting and Social Change 188 (2023): 122271.

[24] Kori, Gururaj S., and Mahabaleshwar S. Kakkasageri. "Classification and regression tree (CART) based resource allocation scheme for wireless sensor networks." Computer Communications 197 (2023): 242-254.

[25] Zhang, Huan, Liangxiao Jiang, and Geoffrey I. Webb. "Rigorous non-disjoint discretization for naive Bayes." Pattern Recognition 140 (2023): 109554.

[26] Moon, Jihoon, Seungwon Jung, Jehyeok Rew, Seungmin Rho, and Eenjun Hwang. "Combination of short-term load forecasting models based on a stacking ensemble approach." Energy and Buildings 216 (2020): 109921.

[27] Marcot, Bruce G., and Anca M. Hanea. "What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis?." Computational Statistics 36, no. 3 (2021): 2009-2031.

[28] Kumar, R. Lokesh, Jagadeesh Kakarla, B. Venkateswarlu Isunuri, and Munesh Singh. "Multi-class brain tumor classification using residual network and global average pooling." Multimedia Tools and Applications 80, no. 9 (2021): 13429-13438.

[29] Chicco, Davide, and Giuseppe Jurman. "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation." BMC genomics 21 (2020): 1-13.

[30] Gurian, Elisa, Alessia Di Silvestre, Elisa Mitri, Devis Pascut, Claudio Tiribelli, Mauro Giuffrè, Lory Saveria Crocè, Valter Sergo, and Alois Bonifacio. "Repeated double cross-validation applied to the PCA-LDA classification of SERS spectra: a case study with serum samples from hepatocellular carcinoma patients." Analytical and bioanalytical chemistry 413 (2021): 1303-1312.

[31] Hou, Yimin, Tao Chen, Xiangmin Lun, and Fang Wang. "A novel method for classification of multi-class motor imagery tasks based on feature fusion." Neuroscience Research 176 (2022): 40-48.

[32] Kanazawa, Naoko, Akari Shigemi, Nao Amadatsu, Kotaro Arimura, Shohei Shimono, Kazutaka Oda, Victor Tuan Giam Chuang, Kazuaki Matsumoto, Hideki Kawamura, and Hideyuki Terazono. "A cohort study of the risk factors and the target AUC to avoid vancomycin-associated acute kidney injury in pediatric patients." *Journal of Infection and Chemotherapy* 30, no. 4 (2024): 323-328.

[33] Guo, Jiongchao, Yating Huang, Lamei Pang, Yuan Zhou, Jingjing Yuan, Bingfeng Zhou, and Minmin Fu. "Association of systemic inflammatory response index with ST segment elevation myocardial infarction and degree of coronary stenosis: a cross-sectional study." *BMC Cardiovascular Disorders* 24, no. 1 (2024): 98.