Industrial Engineering Journal

ISSN: 0970-2555

Volume : 54, Issue 2, February : 2025

A MACHINE LEARNING FRAMEWORK FOR EARLY-STAGE DETECTION OF AUTISM SPECTRUM DISORDERS

[1]DR.MS.KEERTHIGUTTIKONDA,[2]PALLAVARAM SHYAM SUNDAR LAL, [3]REDDI SAI GAYATRIKARTHIKEYA NAIDU,[4]MOHAMMAD BEE BI AYESHA, [5]POTLI LAKSHMI PATHI REDDY,[2,3,4,] 2,3,4,5- IV-B. Tech CSE Students Department of Computer Science and Engineering, Seshadri Rao Gudlavalleru Engineering College (An Autonomous Institute with Permanent Affiliation to JNTUK, Kakinada), Seshadri Rao Knowledge Village, Gudlavalleru521356, Andhra Pradesh, India.

Abstract: The enquiry develops an effective sleuthing system for "Autism Spectrum dieses(ASD)" during early level habituate Machine Learning methodologies despite challenge in full removal of the diseases while attempting early interposition for stiffness reduction. The system includes four "Max Abscaler,PowerTransformer,QuantileTransformer,NormalizerandFeature Scaling(FS)" that analyze their effects on "four standard ASD datasets" from diverse "age gatherings (Babies, Teenagers, Youngsters, and grownup). " "The study algorithmic rule (Ada Boost, K-Nearest neighbor, Decision Tree, Random Forest, Gaussian Naïve Bayes, support vector machine, Logistic Regression, Linear Discriminant Analysis)" are use on let in descale datasets. The organization results are assessed employ numerous actual mensuration, distinguishing which classifiers and part purpose operation work substantially with each age grouping of participant. The testing results show striking precision ascent, with the "Voting classifier anticipating the ASD nearly elevated accuracy of sister and toddler", while it accomplishes the most elevated accuracy for teenagers and adult. The challenge involves acomprehensive examination of the grandness of each element employ four component determination strategies. It suggests that the constituent research can serve hospital therapy experts in the production of conclusion-making through ASD viewing and suggests the implication of raise ML technique for the prediction of ASD across various geezer hood group and socioeconomic backgrounds. The urge structure exhibits ideal consequence in direct contrast with current techniques for other ASD screening. The aim scheme should work on the heartiness and accuracy of ASD recognition; a troupe technique "Adaboost and Random Forest with voting classifier" was perform, execute an uncommon "100% accuracy".

"Index terms **-**"Autism spectrum condition, artificial intelligence-based learning, categorization, data normalization, and attribute selection method."

INTRODUCTION

"Autism Spectrum Disorder (ASD): It is a neuro developmental disease" draw by abnormal mental health that starts in early life, influencing a singular's societal communication and relational capacities. [1], [2]. "Autism Spectrum Disorder (ASD)" is name by "restricted by supernumerary standard of conduct", with the terminus 'range' implying an expansive cluster of side effects and earnestness stage. Albeit a sturdy therapy for ASD stays subtle, early mediation and meet clinical consideration can significantly upgrade a shaver's advance by sharpen on reinforcing ways of behaving and relational abilities [6], [7], [8]. The recognize cogent evidence and finding of ASD stay examination and complex, count on traditional conduct research. Chemical dissymmetry is oftentimes analyzed at two twelvemonths previous, in maliciousness of the fact that it might possibly be analyze later on, dependent upon its seriousness. A setting of helpful methodologies be to efficiently distinguish ASD. These symptomatic strategies are not primarily employed in that body of judgement, where there is a significant risk of prepare "ASD."

The investigator in [12] acquaint brief agenda material that small fry at unlike ages could comprehend across babyhood, childhood, teenagerhood, and adulthood. Consequently, the Creator in [13] fostered

the ASDTests portable application theoretical account for truehearted ASD distinguishing proof, using unlike poll reappraisal, include “Q-Visit and AQ-10 procedures”. Their squad developed an open-germ dataset which included versatile package data and they incite it tothe "The free Machine Learning Repository of the University of California, Irvine (UCI) and the Kaggle platform serve as valuable resources for further research in this field." Different forms of simple machine learning research analyze Mental "Autism Spectrum Disorder (ASD)" along with other circumstances through prediction method acting including tenderness failure and stroke diagnosis and disease valuation [14], [15], [16]. The creator in [17] analyze the element of ASD utilize "Dominion-establish Machine Learning (RML)" methods and dwell out that “RML ascent the truth” of grouping models. "The study in [18] utilizes the Iterative Dichotomizer and Random Forest (RF), (ID3) algorithms to develop a predictive model for kids, newly diagnosed individuals, and grown-up person." The Jehovah in [19] inaugurate an apt judgement device that organize “ADI-R and ADOS ML systems”, utilizing a few qualities encode operation to treat information deficiency, non-one-dimensionality, and irregularity challenge. A result workplace by the Lord in [13] delineates an element to-division and component to-high spot connective respect utilizing genial figuring and sway out "Support Vector Machine (SVM),Logistic Regression (LR) and Decision Tree(DT) " as classifier for the conclusion and guess of ASD [17]. Furthermore, the authors in [20] investigated a conventionally developed (CD) group (N = 19) and individuals with ASD (N = 11), using feature selection based on correlation to determine key attributes." In 2015, the creators in [21] canvas nipper with Mental "Autism Spectrum Disorder (ASD) and typically developing (CD)" untried multitude, pick out 15 preschool ASDs dependent simply upon seven attributes. Besides, they demo that bunch interrogatory may capably inspect unpredictable lesson to cipher ASD total and changeability. "The study in [22] evaluated the classificationaccuracyLogistic Regression (LR),Support vector Machine(SVM),K-Nearest Neighbor(KNN),ClassificationandRegressionTrees (CART), LinearDiscriminant (LD)and Navie Bayes (NB) for predicting adult Autism Spectrum Disorder (ASD)."

LITERATURE SURVEY

We constructed "Data sets of ASD" from infants, children, adolescents, and adults in this research [1], using some component determination methods. "Thus, variouscategorizations were used on these datasets and their execution was compared using measures like prediction accuracy, Auroc,F1 score and the kappa statistics. We also measured the carrying out of discrete classifiers by a non-parametrical statistic significance evaluation. In datasets for infants, adolescents,adults and children. The Support Vector Machine (SVM) produces good results.The RIPPER-based infant subsetwith an accuracy of 97.82%, the kid subset with an accuracy of 99.61% with the Correlation-based Feature Selection (CFS) and Boruta CFS intersect (BIC) method, the Boruta-based adolescent subset an accuracy of 95.87%, and CFS-based adult subset with 96.82% accuracy. Therefore, the application of the Shapley Additive Explanations (SHAP) approach on variouscharacteristic subsets with multiple datasets and obtained optimal unambiguous and classifiedcharacteristicsupon the analysis [1].”."Ongoing examinations concerning the “job of stomach microbiota in wellbeing and dieses” have used 16S quality sequencing of waste examples. Dysbiotic stomach microbiota has been distinguished in "Autism Spectrum Disorder (ASD)", “Neurodevelopmental circumstances” are distinguished by the results of the gastrointestinal aspect. [2]. By the by, regardless of the extensive number of explorations, distinguishing a trademark dysbiotic profile in ASD patients stays testing [3], [4], [5]. The irregularities among this examination emerge from specialized perspectives (i.e., trial systems) and outer factors (i.e., dietary practices). This exploration presents the assortment of 959examples from eight accessible activities, including 540 people with "Autism Spectrum Disorder (ASD) and 419 Healthy Controls (HC)", subsequently alleviating the detailed inclination across cramming. In the same manner, we made use of“Machine Learning (ML) philosophy to stimulatethe point of reference” fit for recognizing “HC and ASD”. We assessed and refined “three algorithms: Gradient Boosting ,Random Forest and

Support vector Machine". Only three algorithms confirmed the significance out of five distinct genera, which includes Alloprevotella and Pasteurella. Also, our discoveries manifest“machine learning algorithms” might observe pervasive ordered characteristics by investigating datasets got from countries with inert frustrating factors. The essential qualities saw in chemical imbalance incorporate lessened relational abilities, hindered connection, and social adjustments. The reasons for these peculiarities can be inspected by an investigation of their visual tactile handling. This exploration uses picture boosts to break down kids' way of behaving by looking at their “visual consideration designs. [3, 4, 5, 9]” The estimation of (FEGP)" has been recommended to notice the look directions of a youngster, examine eye stare qualities, and assess the distinctions in visual discernment between a mentally unbalanced kid and a neurotypical kid. This technique supports perceiving the visual social contrasts in mentally unbalanced youngsters utilizing a presentation level pointer, representation, and surmisings, which might be used to change their learning projects to line up with their companions." Numerous attempts have been made to detect and evaluate neurological disorders manifested through observable symptoms, such as hand tremors. One such condition, multiple sclerosis, can be partially assessed by examining the intensity of hand tremors [5]. This study presents a system developed for capturing and analyzing digital signals in the Spirograph standard test. The development of both hardware and software for this device is detailed, aiming to facilitate the Spirograph test, record signals, transfer them to a computer with dedicated software, and analyze them using feature extraction and classification algorithms. “PowerRange Examination is put forward as an element separated in the product, as it clarifies the impact of every recurrence part on the all-outdevelopment of the hand. The chaoticity stage of the signs is indicated by multiple constraints, such as the largest Lyapunov type and the suggest value of the Lyapunov variety, as determined by the Power Range Examination. Signal intricacy is portrayed by its inserting aspect and delay, which altogether structure a surmised list window for occasional sign reproduction. The delay is connected with the testing rate and the state of the sign. Signals are viewed as examples in highlight space and are ordered by a prepared feedforward brain organization. [16, 20] The characterization task performs as an interactive procedure which evaluates subject enrollments for solid and unhealthy classes while the physicist sets corresponding treatments. This paper shows the way that intricate elements, explicitly turbulent highlights, can really address the unique way of behaving of signs and May be implemented to differentiate between subjects who experience hand tremors and those who do not.“Autism spectrum disorder (ASD)” is a diverse and moderate “neurodevelopmental problem”. Numerous ongoing procedures utilize utilitarian “attractive reverberation imaging (fMRI) to distinguish autism spectrum disorder (ASD) utilizing a limited dataset, yielding brilliant exactness yet bringing about insufficient speculation [3], [4], [5]. " To address this limitation had improve the efficiency of automatic Autism Spectrum Disorder (ASD) detection, this study introduces an ASD classification approach utilizing functional connectivity features extracted from resting-state fMRI data. The initiated method integrates two frequently used brain atlases, Craddock 200 (CC200) and Automated Anatomical Labeling (AAL), along with two less frequently employed atlases, Bootstrap Analysis of Stable Clusters (BASC) and Power. A Deep Neural Network (DNN) classifier is utilized for classification. Exploratory outcomeshows that the initiatedmethodovercome existing approaches in accuracy, achieving an 88% of mean accuracy, whereas state-of-the-art models range between 67% and 85%. Furthermore, the model achieved a sensitivity of 90%, an F1-score of 87%, and The area under receiver operating characteristic curve (AUC) of 96%. A comparative examination of different scoring methodologies highlights the BASC atlas as the most effective in distinguishing ASD from control groups.

METHODOLOGY

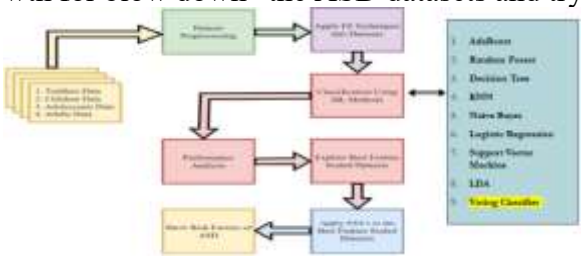
i) Proposed Work:

The advise framework provides an AI bod for the other recognizable evidence of "Autism Spectrum Disorder (ASD)" by utilizing current algorithms and component grading technique, “let in the Quantile

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 2, February : 2025

Transformer, Normalizer,Power Transformer, and Max AbsScaler", to enhance precision and streamline information. The model comport a "comprehensive evaluation of respective ASD datasets across eld groups", with a focus on the enhancement of basic play constituent and the sweetening of focal point desire. Ultimately, this will raise the truth of the demonstrative version. Mechanization and the joining of complex preprocessing procedures ferment withcompelling ASD uncovering, highlighting the significance of early mediation for improved results. A mathematical group methodology employ a "Voting Classifier with RandomForest (RF) and AdaBoost" was fulfill to bring on the versatility and accuracy of ASD identification, bringing about an over-the-top 100 percent truth. This gathering technique coordinates the benefits of a few roles model, using the shifted qualities of "RF and Adaboost for elevate prescient dependableness". To empower node testing, a client labor front end might be built with the "Flask system", guaranteeing a smooth and suck in experience.

ii) System Architecture: This task attend to further a proficient prescient good example utilizing a few machine learning strategy to report genial imbalance across different age gather. "Initially, the datasets are collected and preprocessed, which includes handling missing values, encoding categorical features, and applying oversampling techniques. The Mean Value Imputation (MVI) technique handles the missing data. The One-Hot Encoding (OHE) method is used to convert categorical feature values to their respective analyticalforms. Feature selection is then performed using **eight well-established machine learning classification algorithms**, namely **Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), Logistic Regression (LR), Gaussian Naive Bayes (GNB), K-NearestNeighbors(KNN),Decision Tree(DT),RandomForest(RF)andAdaBoost(AB).**"The orderevent of the classifier on a few "element descale ASD datasets" are looked at, recognizing the best enactment form near and the ideal factor determination method acting for each ASD dataset. Lead to the examination, the gamble broker for ASD is visualize, and the master gamey spot are center along in scene of "their grandness economical time value apply four unmistakable FSTs: GRAE, RFE, CAE, and IGAE (allude to the far tactual sensation methodology in Table 4)". Anatomy 1 represents the purport research growth for blow down "the ASD datasets and try the gamble divisor near powerful in Detection of ASD".



"Fig 1 The Proposed architecture"

iii) collection of Dataset: The module entails the layering and legal opinion of assorted datasets that come to the cover version of ASD across trenchant socioeconomic grouping look on age. It is likely that it will embrace responsibilities such as verifying the social system of the fact, remove knowledge factor, and remove melodic piece of music of information from the dataset.

1." Data of Adult Screening"The Adult Cover dataset encompass data connected to grown-ups and is mainly intended for evaluating Autism Spectrum Disorder (ASD) in individuals past puberty [3], [4], [5]. It might enfold traits like personal doings standards, symmetry capacity, and other appropriate "feature specified to adults for careful ASD appraisal".

Score	A10_Score	...	gender	ethnicity	jundice	austim	contry_of_res	used_app_before	result	age_desc	relation	Class/ASD
0	0	...	f	White-European	no	no	United States	no	6.0	18 and more	Self	NO
0	1	...	m	Latino	no	yes	Brazil	no	5.0	18 and more	Self	NO
1	1	...	m	Latino	yes	yes	Spain	no	8.0	18 and more	Parent	YES
0	1	...	f	White-European	no	yes	United States	no	6.0	18 and more	Self	NO
0	0	...	f	?	no	no	Egypt	no	2.0	18 and more	?	NO

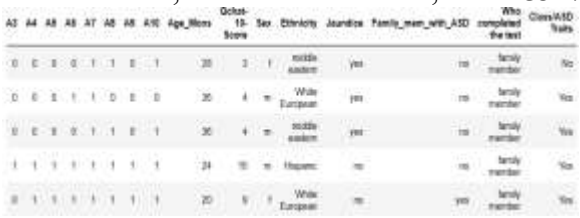
"Fig 2 Dataset of Adult"

Industrial Engineering Journal

ISSN: 0970-2555

Volume : 54, Issue 2, February : 2025

2. "Data of Toddler": -"The toddler dataset focus on the motley bag and testing of data" bear on to short yearling, ofttimes ripen somewhere in the clasp of "one and three years". The purpose of this dataset is to make up one's nous the other signs of ASD, with a special accent on the formative achievements, social associations, "and conveyingability that are similar to this former age group".



"Fig 3 Dataset of Toddler"

3"Adolescent Data": - "The Adolescent dataset" is in all probability collected to look at ASD in the great unwashed get on 12 to 18 year. It might wrap component part that describe the peculiar trouble and character tie to ASD in immaturity, recall alteration for societal way of behaving, communication power, and early relevant variables.

"Fig 4 Dataset of Adolescent"

4. "Data of children": The DataSet which belongs to child data incorporates A Brobdingnagian array of geezer cowling, span from the plastic twelvemonth to pre-puberty. It is seemingly intended to ascertain "ASD-tie in" attributes that are appropriate for child, hold additive such as shaping kid, societal connective, and transmit capacity that are essential to this incision "[3], [4],[5]".

"Fig 5 The Children Dataset"

iv)Processing of the Data:"The Data Processing" is changing over primitive selective information into usable datum for attempt. Datum scientists usually take part in entropy manipulation, enveloping the potpourri, tie, purifying, ratification, test, and change of Data into understandable configurations, like charts or paper. Data Point processing can be led through three "strategies: mechanical, electronic, and manual". The goal is to raise the worth of data and smooth out pilotage. This permits ventures to elevate their activity and fulfill quick essential option. Robotized information handling advancements, let in PC programming writing computer programs, are urgent in this specific situation. It might convert across-the-board datasets, particularly great information, into critical experiences for quality administration and direction.

v) Feature selection: The identification cognitive functioning of near predictable relevant features which are non-superfluous represents feature excerption method. Established process should bound dataset extent because datum book and diversity express on to grow. Constituent pickax forge to amend prognosticative good example feasibility while it minifies the monstrance costs at the same time.Feature choice, a all-important part of percentage designing, admit Agenise the master highlights for input into "machine learning algorithms". The summons of reducing selective information factors look on feature choice methods that extinguish redundant feature of speech to optimize amour between

component for find out system of rules of principle. The essential benefits of restrained component feature article excerption of fourth attribute, instead of brook the motorcar learning model to severally decide the main highlights.

vi) Algorithms: Adaptive Boosting, or AdaBoost, is an approach in machine learning." that further build up order precision by incorporate legion canonical theoretical account. The interaction set out with a principal model, "for example, a solitary storey decisiveness tree", and hence fix supererogatory theoretical account iteratively, doling out more notable grandness to the "entropy focuses" misclassified by the first models. AdaBoost amalgamates these modelling to ramp up a lusty outfit fit for exact expectations, so raise boot batting order fraud detection in your task by arrive at from the blunders of other models and expanding loosely viability.

```
from sklearn.ensemble import AdaBoostClassifier

# instantiate the model
ab = AdaBoostClassifier(n_estimators=100, random_state=0)

# fit the model
ab.fit(X_train, y_train)

y_pred = ab.predict(X_test)
y_prob = ab.predict_proba(X_test)

ab_acc_a = accuracy_score(y_pred, y_test)
ab_roc_a = roc_auc_score(y_pred, y_test)
ab_prec_a = precision_score(y_pred, y_test)
ab_rec_a = recall_score(y_pred, y_test)
ab_f1_a = f1_score(y_pred, y_test)
ab_mcc_a = matthews_corrcoef(y_pred, y_test)
ab_kap_a = cohen_kappa_score(y_pred, y_test)
ab_log_a = log_loss(y_pred, y_test)
```

"Fig 6 Adaboost"

"Random Forest is an ensemble scholarship" operation that ordain various "decision trees" to give prognosis. It exploits via groom a group of "decisiveness trees". The testing process admit average forecasting made from different irregular subsets of selective selective information. This outfit scheme far originate accuracy, mitigates overfitting, and convey red-blooded executing for order and retrogression coating [42].

```
from sklearn.ensemble import RandomForestClassifier

# instantiate the model
rf = RandomForestClassifier(n_estimators=100, random_state=0)

# fit the model
rf.fit(X_train, y_train)

y_pred = rf.predict(X_test)
y_prob = rf.predict_proba(X_test)

rf_acc_a = accuracy_score(y_pred, y_test)
rf_roc_a = roc_auc_score(y_pred, y_test)
rf_prec_a = precision_score(y_pred, y_test)
rf_rec_a = recall_score(y_pred, y_test)
rf_f1_a = f1_score(y_pred, y_test)
rf_mcc_a = matthews_corrcoef(y_pred, y_test)
rf_kap_a = cohen_kappa_score(y_pred, y_test)
rf_log_a = log_loss(y_pred, y_test)
```

"Fig 7Random forest"

Inside the progressive theoretical history of "A Decision Tree" a testing operation exist at the abode knob which upside to analyze terminal result split out into unlike outcomes and conclude leafage lymph node define classification. The graphical display of temporal blueprint follows easily through "Decision Tree Diagram." The poser testify interpretability because it reveals all important ASD prediction epitope by name basic characteristics.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 2, February : 2025

```
from sklearn.tree import DecisionTreeClassifier

# instantiate the model
tree = DecisionTreeClassifier(max_depth=30)

# fit the model
tree.fit(X_train, y_train)

y_pred = tree.predict(X_test)
y_prob = tree.predict_proba(X_test)

dt_acc_a = accuracy_score(y_pred, y_test)
dt_roc_a = roc_auc_score(y_pred, y_test)
dt_prec_a = precision_score(y_pred, y_test)
dt_rec_a = recall_score(y_pred, y_test)
dt_f1_a = f1_score(y_pred, y_test)
dt_mcc_a = matthews_corrcoef(y_pred, y_test)
dt_kap_a = cohen_kappa_score(y_pred, y_test)
dt_log_a = log_loss(y_pred, y_test)
```

"Fig 8 Decision trees"

"K-Nearest Neighbors is a non-parametric procedure" that regularize a data point of stake as per the prevalent class of its "k-nearest neighbors in the component distance". "KNN is knock-down" at pick out invention in data without surmise a specific utilitarian construction. It can mark nearby human relationship inside "The datasets of ASD that may not be free overall [12, 13]".

```
from sklearn.neighbors import KNeighborsClassifier
#from sklearn.pipeline import Pipeline

# instantiate the model
knn = KNeighborsClassifier(n_neighbors=3)

# fit the model
knn.fit(X_train,y_train)

y_pred = knn.predict(X_test)
y_prob = knn.predict_proba(X_test)

knn_acc_a = accuracy_score(y_pred, y_test)
knn_roc_a = roc_auc_score(y_pred, y_test)
knn_prec_a = precision_score(y_pred, y_test)
knn_rec_a = recall_score(y_pred, y_test)
knn_f1_a = f1_score(y_pred, y_test)
knn_mcc_a = matthews_corrcoef(y_pred, y_test)
knn_kap_a = cohen_kappa_score(y_pred, y_test)
knn_log_a = log_loss(y_pred, y_test)
```

"Fig 9KNN:

Naïve Bayes : It is a derived from Bayes' theorem, based on the speculation of independent characteristics. It is computationally adequate and performs well with more regulated datasets.Its effortlessness and stride hand over it agree for the profound assessment of "ASD information".

```
from sklearn.naive_bayes import GaussianNB
#from sklearn.pipeline import Pipeline

# instantiate the model
nb=  GaussianNB()

# fit the model
nb.fit(X_train,y_train)

y_pred = nb.predict(X_test)
y_prob = nb.predict_proba(X_test)

nb_acc_a = accuracy_score(y_pred, y_test)
nb_roc_a = roc_auc_score(y_pred, y_test)
nb_prec_a = precision_score(y_pred, y_test)
nb_rec_a = recall_score(y_pred, y_test)
nb_f1_a = f1_score(y_pred, y_test)
nb_mcc_a = matthews_corrcoef(y_pred, y_test)
nb_kap_a = cohen_kappa_score(y_pred, y_test)
nb_log_a = log_loss(y_pred, y_test)
```

"Fig 10Naïve bayes"

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 2, February : 2025

Logistic Regression is a linear classification model.used for parallel grouping, which approximate the likeliness of an instance hasten a place with a picky category through the strategic capability. “Logistic Regression is interpretable and explains the kinship amongst character and the likeliness of ASD”. It capabilities as a basic world view for parallel characterization under takings.

```
# Logistic Regression model
from sklearn.linear_model import LogisticRegression
#from sklearn.pipeline import Pipeline

# instantiate the model
log = LogisticRegression()

# fit the model
log.fit(X_train,y_train)

y_pred = log.predict(X_test)
y_prob = log.predict_proba(X_test)

lr_acc_a = accuracy_score(y_pred, y_test)
lr_roc_a = roc_auc_score(y_pred, y_test)
lr_prec_a = precision_score(y_pred, y_test)
lr_rec_a = recall_score(y_pred, y_test)
lr_f1_a = f1_score(y_pred, y_test)
lr_mcc_a = matthews_corrcoef(y_pred, y_test)
lr_kap_a = cohen_kappa_score(y_pred, y_test)
lr_log_a = log_loss(y_pred, y_test)
```

“Fig 11Logistic regression”

“Support Vector Machines:The regulated eruditeness fabric that recognizes the hyper plane which ideally recognizes course in a high-layered space”. “Support Vector Machines (SVM)” are capable at overseeing unpredictable choice shortcut points. It can pick out “non-unmediated family relationship in ASD datasets”, consequently advance ordering“accuracy[12,13]”.

```
from sklearn.svm import SVC
svc = SVC()

# fitting the model for grid search
svc.fit(X_train, y_train)

y_pred = svc.predict(X_test)
#y_prob = svc.predict_proba(X_test)

svc_acc_a = accuracy_score(y_pred, y_test)
svc_roc_a = roc_auc_score(y_pred, y_test)
svc_prec_a = precision_score(y_pred, y_test)
svc_rec_a = recall_score(y_pred, y_test)
svc_f1_a = f1_score(y_pred, y_test)
svc_mcc_a = matthews_corrcoef(y_pred, y_test)
svc_kap_a = cohen_kappa_score(y_pred, y_test)
svc_log_a = log_loss(y_pred, y_test)
```

“Fig 12SVM”

Linear Discriminant Analysisis a dimensionality decrease and enactment procedure that know uncoiled mixes of highlights that ideally recognize across course. [23, 26] "Linear Discriminant Analysis (LDA)" is herculean for dimensionality decrease and underlining accredit attributes. It can far develop interpretability and helper with recognizing fundamental standards in ASD identification.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 2, February : 2025

```
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis

clf = LinearDiscriminantAnalysis()

# fitting the model for grid search
clf.fit(X_train, y_train)

y_pred = clf.predict(X_test)
#y_prob = svc.predict_proba(X_test)

lda_acc_a = accuracy_score(y_pred, y_test)
lda_roc_a = roc_auc_score(y_pred, y_test)
lda_prec_a = precision_score(y_pred, y_test)
lda_rec_a = recall_score(y_pred, y_test)
lda_f1_a = f1_score(y_pred, y_test)
lda_mcc_a = matthews_corrcoef(y_pred, y_test)
lda_kap_a = cohen_kappa_score(y_pred, y_test)
lda_log_a = log_loss(y_pred, y_test)
```

"Fig 13 LDA"

"The **Voting Classifier** is a gathered learning technique" that incorporates various separate classifier, whose expectations are collect to impart an ultimate result. This task uses AdaBoost and Random Forest as the profound classifiers.

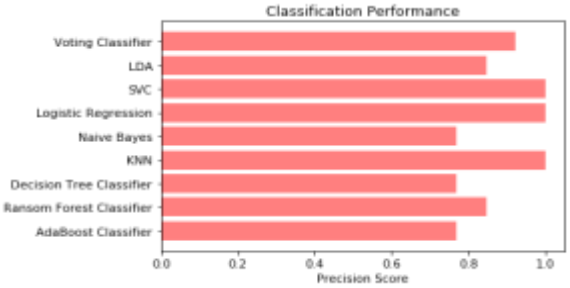
```
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import VotingClassifier
clf1 = AdaBoostClassifier(n_estimators=100, random_state=0)
clf2 = RandomForestClassifier(n_estimators=100, random_state=0)
clf3 = DecisionTreeClassifier(max_depth=30)
eclf1 = VotingClassifier(estimators=[('ab', clf1), ('rf', clf2), ('dt', clf3)], voting='soft')
eclf1.fit(X_train, y_train)
y_pred = eclf1.predict(X_test)

vot_acc_a = accuracy_score(y_pred, y_test)
vot_roc_a = roc_auc_score(y_pred, y_test)
vot_prec_a = precision_score(y_pred, y_test)
vot_rec_a = recall_score(y_pred, y_test)
vot_f1_a = f1_score(y_pred, y_test)
vot_mcc_a = matthews_corrcoef(y_pred, y_test)
vot_kap_a = cohen_kappa_score(y_pred, y_test)
vot_log_a = log_loss(y_pred, y_test)

storeResults('Voting Classifier',vot_acc_a,vot_roc_a,vot_prec_a,vot_rec_a,vot_f1_a,vot_mcc_a,vot_kap_a
```

"Fig 14 Voting classifier"

EXPERIMENTAL RESULTS

"Precision": "Precision valuate the issue of cases" that are exactly ordered among those that have been name as positive. The recipe for operating Precision is subsequently conduct as: "Precision = True positives/ (False positives + True positives) = TP/(FP + TP)"

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$



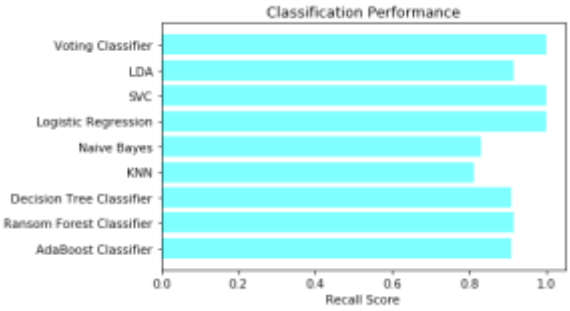
"Fig 15"Precision Performance Comparison

"Recall": The capacity of a manakin to compass all relevant instances of a specific division is valuate by recall, a size of it in system learning. The proportion of on the nose predicted high-pitched-calibre perceptual experience to the total figure of genuine positive, which provides brainstorm into the

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 2, February : 2025

effectiveness of a peculiar version in distinguishing outcome of a specific category.

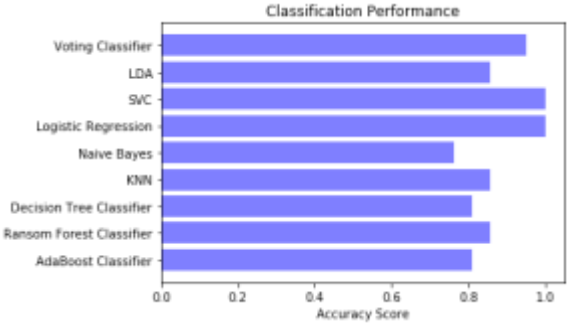
$$Recall = \frac{TP}{TP + FN}$$



"Fig 16 " Recall Performance Comparison"

"Accuracy": The world-wide accuracy of a model's prognosis is appraise by the percent of appropriate anticipation in a characterization challenge.

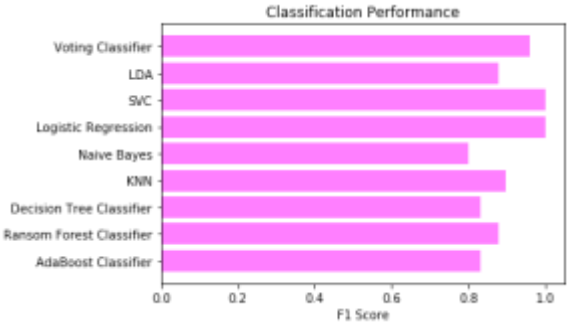
$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$



"Fig 17 graph of Accuracy"

"F1 Score": The F1 Score remains a cost-effective metric befit for imbalanced datasets because it extend dataset information on false positive degree every bit substantially as simulated negatives. The F1 Score mix accuracy time value with thoughtfulness in fiat to function as an efficient single metric

$$\text{F1 Score} = 2 * \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} * 100$$



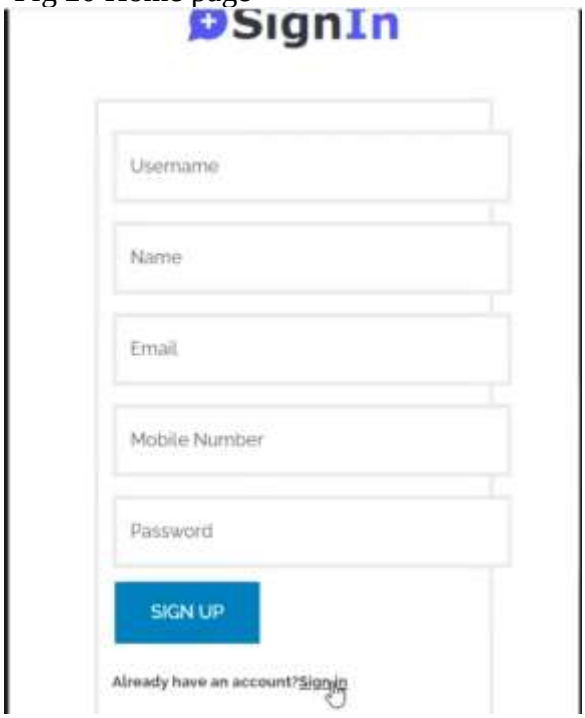
"Fig 18 F1Score"

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 2, February : 2025

ML Model	Accuracy	Precision	Recall	F1-Score
AdaBoost	1.000	1.00	1.000	1.000
Random Forest	1.000	1.00	1.000	1.000
Decision Tree	1.000	1.00	1.000	1.000
KNN	0.945	0.94	0.944	0.922
Naive Bayes	0.979	0.96	0.980	0.970
Logistic Regression	0.993	0.98	1.000	0.990
SVC	0.993	1.00	0.980	0.990
LDA	0.936	0.88	0.936	0.907
Voting Classifier	1.000	1.00	1.000	1.000

"Fig 19 Evolution of the performance"

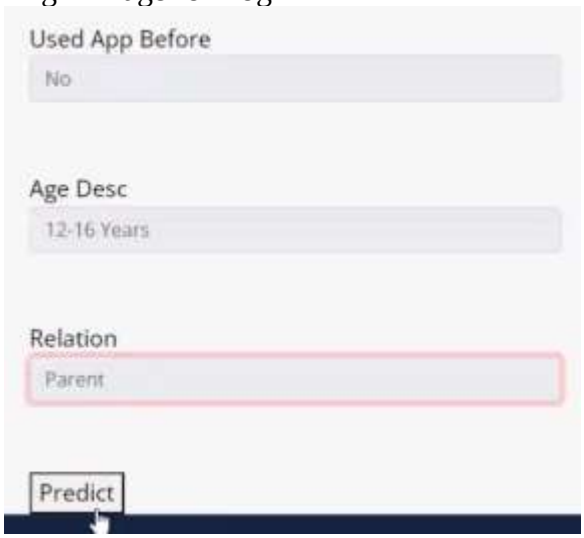
"Fig 20 Home page"



"Fig 21 Page for signin"

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 2, February : 2025

"Fig 22 Page for Login"



"Fig 23Input of the user"

"Fig 24Result predicted for the given input "

CONCLUSION

A smart machine learning scheme was successfully arise for early ASD catching through point A bright political machine learning system has been successfully developed for early ASD detection, utilizing particular scaling and aspect simplification algorithms. The fabric demonstrates warm public presentation across versatile old age group, include toddlers, adolescents, children, and adults. Panoptic evaluations on standard ASD datasets highlight its adaptability and likely clinical significance.This organization describe optimal classification techniques and feature scaling method acting, volunteer a streamlined and effective approach for early ASD espial, which could contribute to timely interventions. The ensemble method acting, integrating Adaboost and Random Forest, has proven extremely efficient in ASD identification, significantly enhancing accuracy.Additionally, the organization is designed for unlined consolidation into a exploiter-friendly interface, provide easy input and testing of feature values, further improving its practicality and usability. By utilize feature selection techniques, the model provides informative ranking of cardinal risk element and crucial aspects decisive to read the complexness of ASD, ultimately aiding in accurate diagnosis.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 2, February : 2025

FUTURESCOPE

This query aimed at increasing the intellectualism of Autism Spectrum Disorder (ASD) and constructing an improved predictive mannequin which could be suitable for person of any age group. The idea is to optimize the spial of ASD as well as other neurological diseases,write future reportmay involve extending thedataset applied to the evaluatio n into a vaster and more heterogeneous subgroup of people who have ASD. The study also suggest the ontogeny of a more universal prognosticativetheoretical model, possibly include new machinelearning methods or improve the current model to enhance the truthfulness and accuracy of ASD identification. The task's future range may admit assessing additional neurodevelopmental disorders and evaluating the proposed theoretical account's effectiveness in detecting and predicting them. Moreover, data aggregation, model sweetening, and potential annex to other neurodevelopmental consideration are also considered central areas for future exploration

REFERENCES

[1] M. Bala, M. H. Ali, M. S. Satu, K. F. Hasan, and M. A. Moni, ‘‘Efficient machine learning models for early-stage detection of autism spectrum disorder,’’ Algorithms, vol. 15, no. 5, p. 166, May 2022.
[2] D. Pietrucci, A. Teofani, M. Milanesi, B. Fosso, L. Putignani, F. Messina, G. Pesole, A. Desideri, and G. Chillemi, ‘‘Machine learning data analysis highlights the role of parasutterella and alloprevotella in autism spectrum disorders,’’ Biomedicines, vol. 10, no. 8, p. 2028, Aug. 2022.
[3] R. Sreedasyam, A. Rao, N. Sachidanandan, N. Sampath, and S. K. Vasudevan, ‘‘Aarya—A kinesthetic companion for children with autism spectrum disorder,’’ J. Intell. Fuzzy Syst., vol. 32, no. 4, pp. 2971–2976, Mar. 2017.
[4] J. Amudha and H. Nandakumar, ‘‘A fuzzy based eye gaze point estimation approach to study the task behavior in autism spectrum disorder,’’ J. Intell. Fuzzy Syst., vol. 35, no. 2, pp. 1459–1469, Aug. 2018.
[5] H. Chahkandi Nejad, O. Khayat, and J. Razjouyan, ‘‘Software development of an intelligent spirography test system for neurological disorder detection and quantification,’’ J. Intell. Fuzzy Syst., vol. 28, no. 5, pp. 2149–2157, Jun. 2015.
[6] F. Z. Subah, K. Deb, P. K. Dhar, and T. Koshiba, ‘‘A deep learning approach to predict autism spectrum disorder using multisite resting-state fMRI,’’ Appl. Sci., vol. 11, no. 8, p. 3636, Apr. 2021.
[7] K.-F. Kollias, C. K. Syriopoulou-Delli, P. Sarigiannidis, and G. F. Fragulis, ‘‘The contribution of machine learning and eye-tracking technology in autism spectrum disorder research: A systematic review,’’ Electronics, vol. 10, no. 23, p. 2982, Nov. 2021.
[8] I. A. Ahmed, E. M. Senan, T. H. Rassem, M. A. H. Ali, H. S. A. Shatnawi, S. M. Alwazer, and M. Alshahrani, ‘‘Eye tracking-based diagnosis and early detection of autism spectrum disorder using machine learning and deep learning techniques,’’ Electronics, vol. 11, no. 4, p. 530, Feb. 2022.
[9] P. Sukumaran and K. Govardhanan, ‘‘Towards voice-based prediction and analysis of emotions in ASD children,’’ J. Intell. Fuzzy Syst., vol. 41, no. 5, pp. 5317–5326, 2021.
[10] S. P. Abirami, G. Kousalya, and R. Karthick, ‘‘Identification and exploration of facial expression in children with ASD in a contact less environment,’’ J. Intell. Fuzzy Syst., vol. 36, no. 3, pp. 2033–2042, Mar. 2019.
[11] M. D. Hossain, M. A. Kabir, A. Anwar, and M. Z. Islam, ‘‘Detecting autism spectrum disorder using machine learning techniques,’’ Health Inf. Sci. Syst., vol. 9, no. 1, pp. 1–13, Dec. 2021.
[12] C. Allison, B. Auyeung, and S. Baron-Cohen, ‘‘Toward brief ‘red flags’ for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls,’’ J. Amer. Acad. Child Adolescent Psychiatry, vol. 51, no. 2, pp. 202–212, 2012.

[13] F. Thabtah, F. Kamalov, and K. Rajab, ‘‘A new computational intelligence approach to detect autistic features for autism screening,’’ Int. J. Med. Inform., vol. 117, pp. 112–124, Sep. 2018.
[14] M. M. Ali, B. K. Paul, K. Ahmed, F. M. Bui, J. M. W. Quinn, and M. A. Moni, ‘‘heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison,’’ Comput. Biol. Med., vol. 136, Sep. 2021, Art. no. 104672.
[15] E. Dritsas and M. Trigka, ‘‘Stroke risk prediction with machine learning techniques,’’ Sensors, vol. 22, no. 13, p. 4670, Jun. 2022
[16] V. Chang, J. Bailey, Q. A. Xu, and Z. Sun, ‘‘Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms,’’ Neural Comput. Appl., early access, pp. 1–17, Mar. 2022.
[17] F. Thabtah, ‘‘Machine learning in autistic spectrum disorder behavioral research: A review and ways forward,’’ Inform. Health Social Care, vol. 44, no. 3, pp. 278–297, 2018.
[18] K. S. Omar, P. Mondal, N. S. Khan, M. R. K. Rizvi, and M. N. Islam, ‘‘A machine learning approach to predict autism spectrum disorder,’’ in Proc. Int. Conf. Electr., Comput. Commun. Eng. (ECCE), Feb. 2019, pp. 1–6.
[19] H. Abbas, F. Garberson, E. Glover, and D. P. Wall, ‘‘Machine learning approach for early detection of autism by combining questionnaire and home video screening,’’ J. Amer. Med. Informat. Assoc., vol. 25, no. 8, pp. 1000–1007, 2018.
[20] K. L. Goh, S. Morris, S. Rosalie, C. Foster, T. Falkmer, and T. Tan, ‘‘Typically developed adults and adults with autism spectrum disorder classification using centre of pressure measurements,’’ in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Mar. 2016, pp. 844–848.
[21] A. Crippa, C. Salvatore, P. Perego, S. Forti, M. Nobile, M. Molteni, and I. Castiglioni, ‘‘Use of machine learning to identify children with autism and their motor abnormalities,’’ J. Autism Develop. Disorders, vol. 45, no. 7, pp. 2146–2156, 2015.
[22] B. Tyagi, R. Mishra, and N. Bajpai, ‘‘Machine learning techniques to predict autism spectrum disorder,’’ in Proc. IEEE Punecon, Jun. 2019, pp. 1–5.
[23] F. Thabtah and D. Peebles, ‘‘A new machine learning model based on induction of rules for autism detection,’’ Health Informat. J., vol. 26, no. 1, pp. 264–286, Mar. 2020.
[24] M. Duda, R. Ma, N. Haber, and D. P. Wall, ‘‘Use of machine learning for behavioral distinction of autism and ADHD,’’ Transl. Psychiatry, vol. 6, no. 2, pp. e732–e732, Feb. 2016.
[25] S. B. Shuvo, J. Ghosh, and A. S. Oyshi, ‘‘A data mining-based approach to predict autism spectrum disorder considering behavioral attributes,’’ in Proc. 10th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT), Jul. 2019, pp. 1–5.
[26] O. Altay and M. Ulas, ‘‘Prediction of the autism spectrum disorder diagnosis with linear discriminant analysis classifier and K-nearest neighbor in children,’’ in Proc. 6th Int. Symp. Digit. Forensic Secur. (ISDFS), Mar. 2018, pp. 1–4.
[27] F. N. Buyukoflaz and A. Ozturk, ‘‘Early autism diagnosis of children with machine learning algorithms,’’ in Proc. 26th Signal Process. Commun. Appl. Conf. (SIU), May 2018, pp. 1–4.
[28] M. F. Misman, A. A. Samah, F. A. Ezudin, H. A. Majid, Z. A. Shah, H. Hashim, and M. F. Harun, ‘‘Classification of adults with autism spectrum disorder using deep neural network,’’ in Proc. 1st Int. Conf. Artif. Intell. Data Sci. (AiDAS), Sep. 2019, pp. 29–34.
[29] S. Huang, N. Cai, P. P. Pacheco, S. Narrandes, Y. Wang, and W. Xu, ‘‘Applications of support vector machine (SVM) learning in cancer genomics,’’ Cancer Genomics Proteomics, vol. 15, no. 1, pp. 41–51, Jan./Feb. 2018.
[30] A. S. Haroon and T. Padma, ‘‘An ensemble classification and binomial cumulative based PCA for diagnosis of Parkinson’s disease and autism spectrum disorder,’’ Int. J. Syst. Assurance Eng. Manage., early access, pp. 1–16, Jul. 2022.
[31] R. Abitha, S. M. Vennila, and I. M. Zaheer, ‘‘Evolutionary multiobjective optimization of artificial neural network for classification of autism spectrum disorder screening,’’ J. Supercomput., vol. 78, no. 9, pp. 11640–11656, Jun. 2022.

[32] M. Alsuliman and H. H. Al-Baity, ‘‘Efficient diagnosis of autism with optimized machine learning models: An experimental analysis on genetic and personal characteristic datasets,’’ Appl. Sci., vol. 12, no. 8, p. 3812, Apr. 2022.
[33] S. P. Kamma, S. Bano, G. L. Niharika, G. S. Chilukuri, and D. Ghanta, ‘‘Cost-effective and efficient detection of autism from screening test data using light gradient boosting machine,’’ in Intelligent Sustainable Systems. Singapore: Springer, pp. 777–789, 2022.
[34] U. Gupta, D. Gupta, and U. Agarwal, ‘‘Analysis of randomization-based approaches for autism spectrum disorder,’’ in Pattern Recognition and Data Analysis with Applications. Singapore: Springer, pp. 701–713, 2022.
[35] T. Akter, M. Shahriare Satu, M. I. Khan, M. H. Ali, S. Uddin, P. Lio, J. M. W. Quinn, and M. A. Moni, ‘‘Machine learning-based models for early-stage detection of autism spectrum disorders,’’ IEEE Access, vol. 7, pp. 166509–166527, 2019.
[36] Kaggle. (2022). Autism Spectrum Disorder Detection Dataset for Toddlers. [Online]. Available: https://www.kaggle.com/fabdelja/autism-screeningfor-toddlers
[37] UCI. (2022). UCI Machine Learning Repository: Autistic Spectrum Disorder Screening Data for Adolescent Data Set. [Online]. Available: https://shorturl.at/fhxCZ
[38] UCI. (2022). UCI Machine Learning Repository: Autism Screening Adult Data Set. [Online]. Available: https://archive.ics.uci.edu/ ml/datasets/Autism+Screening+Adult
[39] UCI. (2022). UCI Machine Learning Repository: Autistic Spectrum Disorder Screening Data for Children Data Set. [Online]. Available: https://shorturl.at/fiwLU
[40] D. Singh and B. Singh, ‘‘Investigating the impact of data normalization on classification performance,’’ Appl. Soft Comput., vol. 97, Dec. 2020, Art. no. 105524.
[41] D. Mease, A. J. Wyner, and A. Buja, ‘‘Boosted classification trees and class probability/quantile estimation,’’ J. Mach. Learn. Res., vol. 8, no. 3, pp. 409–439, 2007.
[42] Q. Wang, W. Cao, J. Guo, J. Ren, Y. Cheng, and D. N. Davis, ‘‘DMP_MI: An effective diabetes mellitus classification algorithm on imbalanced data with missing values,’’ IEEE Access, vol. 7, pp. 102232–102238, 2019.
[43] S. M. M. Hasan, M. A. Mamun, M. P. Uddin, and M. A. Hossain, ‘‘Comparative analysis of classification approaches for heart disease prediction,’’ in Proc. Int. Conf. Comput., Commun., Chem., Mater. Electron. Eng. (ICME), Feb. 2018, pp. 1–4.
[44] D. Ramesh and Y. S. Katheria, ‘‘Ensemble method based predictive model for analyzing disease datasets: A predictive analysis approach,’’ Health Technol., vol. 9, no. 4, pp. 533–545, Aug. 2019.
[45] A. Arabameri and H. R. Pourghasemi, ‘‘Spatial modeling of gully erosion using linear and quadratic discriminant analyses in GIS and R,’’ in Spatial Modeling in GIS and R for Earth and Environmental Sciences. Amsterdam, The Netherlands: Elsevier, pp. 299–321, 2019.