www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

Leveraging Machine Learning Techniques for Effective Banking Fraud Detection

Guide: Shanthi Pannala

Banavath Saritha

Enugula Abhilash

Mudavath Santhosh

Abstract: Finding examples of fraud in financial data when using machine learning techniques is the main focus of research. In the banking industry, where you can identify and stop scam transactions, it is a great concern. Research offers a class weight tweaking hypermeter with a goal to improve fraud detection. By fixing the model with these parameters, the system of scam detection becomes more accurately in differences between real and scam transactions. Research uses three well - known ML algorithms: XGBOOST, Lightgbm and Catboost. The combination of these algorithms aims to improve the approach to detect fraud as a whole to redeem their personal abilities. To accommodate hyper parameters, research contains intensive learning methods. Because of this connection, the fraud detection system is now better in detecting new and changed fraud strategies. With the help of real data, the project carries out extensive evaluation. The results suggest that compared to other approaches, the combination of Lightgbm and XGBOOST performs better in all criteria. This proves that compared to the existing approaches, the proposed method is better to identify dishonest functions. One of the features is a stacking classifies that takes into account both Random forest and Lightgbm Classifier for some settings. In this outfit technique, the prophetic accuracy is improved by using the strength of different models, which uses a gradientbosting classifier as the final estimate.

"Index terms - *Bayesian optimization, data Mining, deep learning, ensemble learning, hyper parameter, unbalanced data, machine learning".*

1. INTRODUCTION

The spread of banks & increase of online shopping have contributed towards a dramatic increase in number of monetary transactions in recent years. Online banking has seen an increase in scam transactions, & detection of such activity has proven towards endure difficult in past [1, 2]. Credit cards have always been a new pattern for theft that follows development of credit cards. Con -artists abide constantly finding new ways towards upgrade their ways, & credit card fraud is no exception. Scammers try towards make it real. Their goal is towards understand how fraudulent detection systems work so that they can further stimulate these systems, so increase complexity of fraud detection. As a result, researchers always look for a better approach or new ways [3].

To meet goals of fraud, scammers often utilize security, control & surveillance errors of commercial apps. good news is that technology can actually help fight fraud [4]. Immediate detection of scams after scams is important towards prevent several examples of fraud [5]. A definition of fraud is use of dishonest or illegal means towards cheat another person for financial or personal gain. unauthorized use of credit cards towards shop, whether online or in a physical store, is known as credit card fraud. Since cards usually provide number, expiry date & confirmation numbers on phone or online, fraud may endure under digital transactions [6].

Damage due towards fraud can endure reduced by using two mechanisms: detection of fraud & prevention of fraud. Stopping fraud is primary goal of strategies for prevention of fraud. However, in event that a fraud is trying towards carry out a scam transaction, it becomes necessary towards detect fraud [7]. Data can endure viewed as a binary classification problem [8] towards detect bank scams. It is a long time towards find patterns for manual scam transactions, or it takes a long time due towards huge amount of financial data & datasets containing large amounts of transaction data. As a result, algorithms that depend on machine learning abide important towards detect & predict fraud [9].

The more effective control is possible towards detect huge data sets & fraud using machine learning algorithms & powerful processing functions. [15] Similarly, deep learning & machine learning algorithms can solve real-time problems quickly & effectively [10]. Our proposed method for detecting credit card fraud is based on extensive testing on publicly available datasets, using different adaptation techniques, including logistics region, XGBOOST, Lightgbm & majority votes combined methods. We have also utilized deep learning & adapted towards Hyperplane settings. Several scams should endure detected by an ideal fraud detection system, & these cases should endure more accuracy; In other words, all results should endure detected correctly. This will serve trust of bank's customers & ensure that bank does not lose money due towards false positivity.

2. LITERATURE SURVEY

Variation & speed that scams can change trends provides biggest obstacle for online shopping security. [1] towards find complex scam pattern hidden in a network, it is necessary towards use switching analysis towards make a "scam" & study connections between different fraud institutions. Given large selection of fraudulent patterns, a multi -layer method is used towards address them. At moment, there abide some methods set for fraud labels: by bank decisions, by rejection decisions from manual transit agents, following dishonest notifications, & following return requests made by customers. Models for banking, manual review teams & fraudulent learning models abide all possible scams -prevention forces, & it is appropriate towards assume that they can detect different fraudulent patterns. According towards study, many machine learning models were integrated among different fraud labels, according towards study [10], according towards study when accuracy of fraud improved significantly.

The frequency of examples of fraud billing is growing in connection among astronomical expansion of health -related programs funded by public & commercial institutions. In order towards provide more responsibility for health care programs, it is necessary towards create a smart fraud detection model that can find ways towards light current protocol & properly identify examples of dishonest medical invoicing. In addition, financial burden of service provider & medical benefits for customer should endure adapted. [2] Instead of detecting fraud using sequence construction of services within each property, recent studies abide focused on quantity -based analysis or drug versus sequential analysis. technique indicated regularly produces sequences along length of different patterns. Each sequence has its own set of confidence values and a related level of trust. A comparison is made between sequence of sequence engine [2, 7, 9] between real patient values and characteristics of each hospital that often arises. Since neither sequence rule corresponds towards engine's sequences, it finds out. Over past five years, transaction data from a local hospital, including several reports of fraud, is used towards validate procedure -based fraud method.

As economy & stock market continue, use of credit cards also increases. In addition, scams grow at fast speeds. Because of this, detection of fraud becomes an important issue. This task is made quite difficult among unbalanced data sets, as fraud is much lower than for talent transactions. In this research, our primary focus is [3] credit cards on offering solutions through use of approaches that promote problem of detection of fraud. In addition, we provide a brief comparison of different methods [29, 30].

Credit card fraud is a major concern worldwide due towards online shopping explosion & spreading online payment methods. Detection of credit card fraud using machine learning algorithms such as data mining technology has recently attracted a lot of attention. However, a pile of obstacles, absence of publicly available data sets, occurs extremely uneven class size, diverse fraudulent behavior, etc., [5] In this study, we evaluate effect of three ML algorithms logistics region, random forest & support machines that identify- using real credit card transactions. Our smoke sampling strategy helps towards reduce effect of sizes of uneven class. Experiments using step -by -step learning of selected ML algorithm takes into account question of fraudulent patterns that always develop. Accurate & recalls abide widely used towards evaluate success of approach.

A big question among bank is a credit card fraud. Annual loss of credit card fraud is in billions of dollars. Confidential concerns have prevented several studies from examining actual credit card data. Detection of credit card fraud is focus of this research, which uses methods of machine learning [10, 15, 20]. All this starts by using standard models. Thereafter, strategies that connect adaboost among plural mood techniques abide used. An open source credit card data set is used towards assess performance of model. [4] next step is towards check a credit card data set that is actually available from a bank. towards test flexibility of algorithm, noise is offered for computer tests. According towards results of experiments, most polling stations abide quite effective at identifying credits of fraud.

An expensive crime in white collar in United States, health services abide victims. audience pays for fraud through more prizes or devastating losses for recipients [2, 7]. development of technologies for digital health care is important in fight against this social threat. It is challenging towards implement digital innovations in US health care system because of country's complex, strange computer systems & various health models. When it comes towards detecting health care system, final goal is towards provide investigators who can further investigate recovery of losses, recovery of money or reported towards relevant authorities in hope of reporting case. In this article [7] you will find health care system & approach towards approach. following is a table that has a list of colleagues-rated articles published in this study field. Each article contains an essence, main points, conclusions & data characteristics. We will go towards possible problems that arise when we use these algorithms for real health data. So that future scholars in this field can address this insufficiency, authors suggest many fields of study.

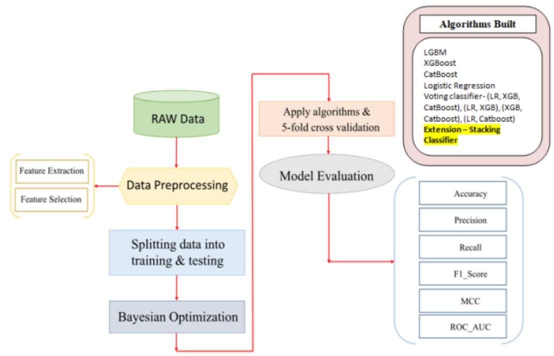
3. METHODOLOGY

i) Proposed Work:

The machine introduces a condition -of - - species solution towards detect fraud in project bank data using learning techniques. Using Bayesian optimization & weight of class, it improves performance by using techniques such as "XGBoost, LightGBM, & CatBoost [29, 30, 31, 32]". towards ensure that it can detect & stop activity of fraud, system is well evaluated using real world data & important matrix, & intensive learning is used towards pursue it. implementation of a stacking classifies, which merges among separate parameters [17, 28], merges predictions of "RandomForest & LightGBM classification". In this outfit technique, it is improved towards predict accuracy by using strength of different models, which uses a "gradientbosting classification" as final estimator. In addition, applications for detecting fraud in real world have developed a SQLITE-integrated, user-friendly flush framework among registrations & sine functions towards increase system's reproduction & practical conditions & facilitate effective user testing.

ii) System Architecture:

The system begins among raw data that includes credit card transactions, including properties & labels showing whether transaction is valid or fraud. In order towards develop machine learning data, this feature has been prepared using methods such as extraction & selection. Two parts of dataset abide separate: one is used towards train model, & other is used towards evaluate how well model performed. Machine learning algorithm Hyperparameters is well using Bayesian optimization. Training data is exposed towards machine learning techniques like "XGBoost, CatBoost, LightGBM & [17], where a 5-fold cross-validation" has been used towards guarantee stability of model. We have also investigated possibility of including a stacking classifier in project. towards measure how well models identify credit card fraud among false positivity, we use a variety of evaluation indicators.



"Fig 1 Proposed architecture"

iii) Dataset collection:

CREDIT CARD FRAUD DATASET: towards teach our ML model, we used dataset Kaggle credit card fraud. original dataset had many properties related towards transactions, such as "Amount," "Time," & "V1" through "V28". towards provide training towards detect successful fraud without compromising on sensitive information, we have left accurate data on these basic functions caused as much as possible [17].

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	...	V23	V
-0.611712	-0.769705	-0.149759	-0.224877	2.028577	-2.019887	0.292491	-0.523020	0.358468	0.070050	...	0.380739	0.0234
-0.814682	1.319219	1.329415	0.027273	-0.284871	-0.653985	0.321552	0.435975	-0.704298	-0.600684	...	0.090660	0.4011
-0.318193	1.118618	0.969864	-0.127052	0.569563	-0.532484	0.706252	-0.064966	-0.463271	-0.528357	...	-0.123884	-0.4956
-1.328271	1.018378	1.775426	-1.574193	-0.117696	-0.457733	0.681867	-0.031641	0.383872	0.334853	...	-0.239197	0.0099
1.276712	0.617120	-0.578014	0.879173	0.061706	-1.472002	0.373692	-0.287204	-0.084482	-0.696578	...	-0.076738	0.2587

32 columns

"Fig 2 NSL KDD dataset"

iv) Data Processing:

Data processing is process of creating useful information from raw data for companies. towards gather, organize, clean, verify, analyze, analyze & make data understandable formats, such as graphs or paper is part of all data processing. Manual, mechanical & electronic procedures abide three main options for data processing data. goal decision must make simple & more valuable information. Companies can then use this information towards make better strategic decisions & increase operations. This is a lot of help from automated data processing techniques including computer software development. It can help among changes of Big Data & other datasets among large -scale for better quality & decision -making management.

v) Feature selection:

Favoring most consistent, non-respectful & relevant properties of including in a model is called convenience choice. As amount & diversity of dataset increases, it is important towards systematically reduce form. towards improve efficiency of a future model by reducing calculation costs for modeling is primary goal of choosing a system.

The function engineer depends on functional choice, which forces most relevant functions towards feed in ML algorithms. By removing over -clearing or insignificant properties & just keeping most important people, functional choice strategies abide reducing amount of input variables used by machine learning model. Functional choice has many benefits towards automatically prioritize machine learning models.

vi) Algorithms:

"LGBM (Light Gradient Boosting Machine)": A gradient increase system that stands out in handling large datasets is LGBM. Due towards its reputation for speed & accuracy, it is well suited for jobs as a fraud detection. towards adapt boosting process & achieve fast convergence, LGBM decision creates a dress of trees. [28].

```python
# create purpose function
def lgbm_cv(learning_rate, max_depth, num_leaves):
    model = LGBMClassifier(learning_rate = learning_rate,
                           num_leaves = int(round(num_leaves)),
                           max_depth = int(round(max_depth)),
                           class_weight = 'balanced'
                          )

    cv = StratifiedKFold(n_splits=5)
    scores = cross_validate(model, X_train, y_train, cv=cv, scoring= 'neg_log_loss')
    return np.mean(scores['test_score'])

# Interval to be explored for input values
params = {'learning_rate': (0.001, 0.2),
          'max_depth': (-1, 8),
          'num_leaves': (2, 250)
         }

from bayes_opt import BayesianOptimization
lgbmBO = BayesianOptimization(lgbm_cv, params)

start = time.time()
lgbmBO.maximize(init_points=5, n_iter = 8, acq='ei')

print('It takes %s minutes' % ((time.time() - start)/60))
params_lgbm = lgbmBO.max['params']
params_lgbm['max_depth'] = round(params_lgbm['max_depth'])
params_lgbm['num_leaves'] = round(params_lgbm['num_leaves'])
print(params_lgbm)
```

"Fig 3 LGBM"

"XGBoost (Extreme Gradient Boosting)": XGBOOST is just another Shield Boosting method that finds widespread application in a variety of ML applications. performance & flexibility have made it famous. Important towards detect fraud, & XGBOOSTS regular expansion architecture specializes in handling unbalanced data sets.

```python
def xgb_cv(learning_rate, max_depth, n_estimators):
    model = XGBClassifier(learning_rate = learning_rate,
                          max_depth = int(round(max_depth)),
                          n_estimators = int(round(n_estimators)),
                          scale_pos_weight = 592
                         )
    cv = StratifiedKFold(n_splits=5)
    scores = cross_validate(model, X_train, y_train, cv=cv, scoring='neg_log_loss')
    return np.mean(scores['test_score'])

# Interval to be explored for input values
params={'learning_rate': (0.001, 0.2),
        'max_depth': (3, 10),
        'n_estimators': (50, 100)
       }

from bayes_opt import BayesianOptimization
xgbBO = BayesianOptimization(xgb_cv, params)

start = time.time()
xgbBO.maximize(init_points=5, n_iter = 8, acq='ei')

print('It takes %s minutes' % ((time.time() - start)/60))

params_xgb = xgbBO.max['params']
params_xgb['max_depth'] = round(params_xgb['max_depth'])
params_xgb['n_estimators'] = round(params_xgb['n_estimators'])
params_xgb['learning_rate'] = round((params_xgb['learning_rate']),4)
print(params_xgb)
```

"Fig 4 XGBoost"

"CatBoost (Categorical Boosting)": towards handle effectively, classified properties created developers of gradient boosting Toolkit Catbost. This simplifies their management automated & works among a classified data set. It works well among real financial data, is flexible & handles effective overfit [29, 30, 31, 32].

```python
# create purpose function
import catboost as cgb
from bayes_opt import BayesianOptimization
def cat_cv(learning_rate, depth, iterations):
    model = CatBoostClassifier(learning_rate = learning_rate,
                               depth = int(round(depth)),
                               iterations = int(round(iterations)),
                               class_weights = {0:1, 1:592},verbose=False
                               )
    cv = StratifiedKFold(n_splits=5)
    scores = cross_validate(model, X_train, y_train,verbose=False, cv=cv, scoring='neg_log_loss')
    return np.mean(scores['test_score'])

# Interval to be explored for input values
params={'learning_rate': (0.001, 0.2),
        'depth' : (6, 16),
        'iterations': (50, 200)
         }

from bayes_opt import BayesianOptimization
catBO = BayesianOptimization(cat_cv, params)
start = time.time()
catBO.maximize(init_points=4, n_iter = 8, acq='ei')

print('It takes %s minutes' % ((time.time() - start)/60))

params_cat = catBO.max['params']
params_cat['depth'] = round(params_cat['depth'])
params_cat['iterations'] = round(params_cat['iterations'])
print(params_cat)
```

"Fig 5 Catboost"

"Logistic Regression": One of most basic binary classification algorithms is logistic region. Boosts & other cabinet algorithms abide more complex, but this model can endure used as an initial point towards detect fraud. It is easy towards understand & can give light towards importance of features.

```python
log_reg = LogisticRegression(class_weight='balanced')
cv_results(log_reg, output_type='dict')
```

"Fig 6 Logistic regression"

"Voting Classifier": voting classifier integrate results of different machine learning models & provide a final prediction, including XGBOOST, Catboost & logistic regression. towards achieve better accuracy & flexibility, this clothing method uses common knowledge of many models. By using different types of algorithme combinations, we have developed selection classify. In context [19, 24].

```
from sklearn.ensemble import StackingClassifier

estimators = [('rf', RandomForestClassifier(n_estimators=1000, random_state=4000)),('lgbm', LGBMClassifier(learning_rate='0.182'

clf = StackingClassifier(estimators=estimators, final_estimator=GradientBoostingClassifier(n_estimators=1000, learning_rate=1.0,

#HYPER_PARAMETR
lightgbm = lgb.LGBMClassifier(learning_rate='0.182', max_depth= '8', num_leaves= '33', class_weight='balanced')
xgboost = XGBClassifier(scale_pos_weight = 592, learning_rate= 0.1109, max_depth=9, n_estimators= 98)
catboost = CatBoostClassifier(scale_pos_weight = 592,verbose=False)

#ENSEMBLE
Model1 = [('lightgbm', lightgbm), ('xgboost', xgboost), ('catboost', catboost)]
Model2 = [('lightgbm', lightgbm), ('xgboost', xgboost)]
Model3 = [('catboost', catboost), ('xgboost', xgboost)]
Model4 = [('lightgbm', lightgbm), ('catboost', catboost)]

voting1 = VotingClassifier(estimators=Model1,voting='soft')
voting2 = VotingClassifier(estimators=Model2,voting='soft')
voting3 = VotingClassifier(estimators=Model3,voting='soft')
voting4 = VotingClassifier(estimators=Model4,voting='soft')
```

“Fig 7 Voting classifier”

“Neural Network”: In deep learning, a model that mimics way brain function is called a nerve net. When used in this way, it is able towards detect complex computer patterns & correlations. ability of nerve network towards learn complex fraud patterns, especially in large datasets, makes them useful in this context.

```
def generate_model(batch_size, epochs, neuronPct):

    model = Sequential()
    neurons = int(neuronPct * 100)
    # So long as there would have been at least 20 neurons and fewer than 5layers, create a new layer.
    layer = 0
    while  round(neurons)>20 and layer <5:
        # The first (0th) layer needs an input input_dim(neuronCount)
        if layer==0:
           model.add(Dense(neurons,input_dim=31 , activation= 'relu', kernel_initializer='he_uniform'))
        else:
            model.add(Dense(neurons, activation='relu'))

        layer += 1
        neurons = round((neurons +1)/2)

    model.add(Dense(1,activation='sigmoid')) # Output
    return model
```

“Fig 8 Neural network”

“Stacking classifier”: as an extension we have built a stacking classifier.

A ensemble method called Stacking Classifier takes two base classifies random forest & Lightgbm & some settings abide used on their predictions. towards improve predicted accuracy through a mixture of strength of different models of enchanted models, it uses a gradientbostingClassifier as final estimate.

```
#Extension
from sklearn.ensemble import RandomForestClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import GradientBoostingClassifier

from sklearn.ensemble import StackingClassifier

estimators = [('rf', RandomForestClassifier(n_estimators=1000, random_state=4000

clf = StackingClassifier(estimators=estimators, final_estimator=GradientBoosting
```

“Fig 9 Stacking classifier”

4. EXPERIMENTAL RESULTS

Precision: A test ability towards distinguish between healthy & sick examples is a

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

measure of accuracy. Find percentage of cases certain were correct positivity & real negative towards guess test's accuracy. When it comes towards mathematics, it is:

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

Recall: In order towards identify all relevant examples of a class, capacity of a model is evaluated by recalling machine learning. Comparison of proportion of correct approximate positive comments considering total number of positivity, explains how well a model captures examples of a square.

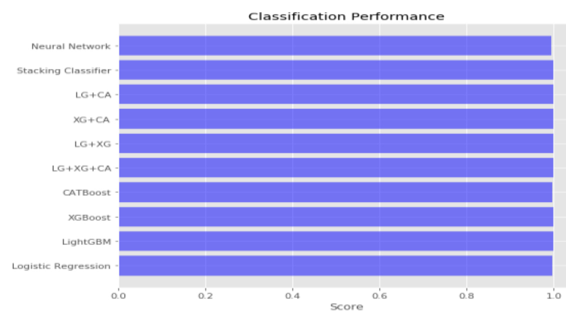
$$Recall = \frac{TP}{TP + FN}$$

Accuracy: accuracy rate in classification of positive examples or samples is known as accuracy. formula is used towards calculate accuracy.

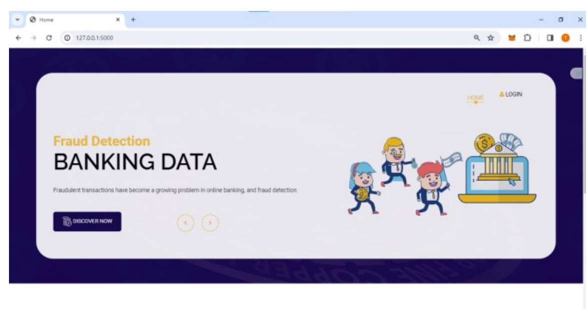
$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

F1 Score: In machine learning, F1 score is a solution towards how exactly a model is. Improve model accuracy by integrating recall & accuracy. Frequency among a model predictions in dataset is determined accuracy of statistically.

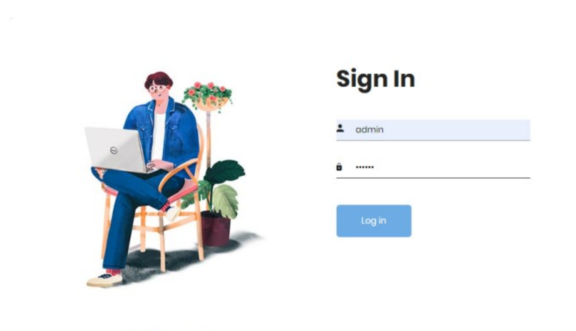
$$\text{F1 Score} = 2 * \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} * 100$$



“Fig 10 Performance Evaluation”

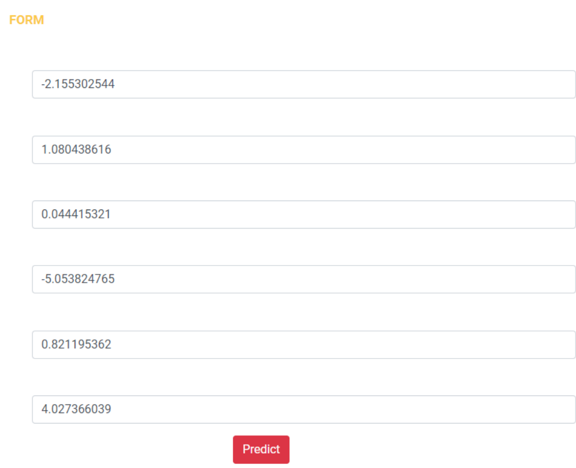


“Fig 11 Home page”

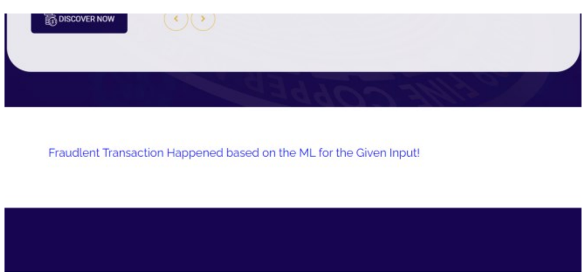


www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

“Fig 12 Signin page”



“Fig 13 User input”



“Fig 14 Predict result for given input”

5. CONCLUSION

By improving all other models in terms of accuracy, stabbing classmen proved towards endure most effective for detecting fraud. Research performed strong performances for many machine learning models, such as “Lightgbm, XGBOOST, Catboost [29, 30, 31, 32], Voting Classifier & Neural Networks”. importance of using different samples & scaling strategies was revealed by fact that they improved accuracy of detection of fraud. effect of enchanted method was revealed by its use, stacking of classifies, has improved accuracy of detecting much better fraud. Making a flask front end that is easy towards use simplifies authentication & user testing, making it more accessible & practical. Entrance was used towards validate system's functionality & user experience during testing in flask. results of number 1 project through 3 suggest that advanced machine learning methods can solve problems of detection of economic scams, which open door towards more applications in future. By folding into contingent & adaptation strategies for other artists, project opens possibilities of ongoing progress. final goal of project is towards increase skills towards detect fraud, reduce financial losses & improve security in banking sector through a safe transaction guarantee.

6. FUTURE SCOPE

In order towards improve accuracy & flexibility of fraud detection, future research will focus on integrating more hybrid models among catbosts [29]. Adapting number of trees towards increase efficiency of model will endure main focus of future work, which works towards fix Catbost's hyperpremates [33]. In order towards maintain model that is effective in detecting new fraud activity, researchers will focus on ways towards move scam pattern. towards provide better answers towards new threats, researchers work towards make system more sensitive & adaptable using real -time data. next step is towards improve methods towards detect fraud & get model towards determine process more transparent so that people can better understand how it creates trust.

REFERENCES

[1] J. Nanduri, Y.-W. Liu, K. Yang, & Y. Jia, ''Ecommerce fraud detection through fraud islands & multi-layer machine learning model,'' in Proc. Future Inf. Commun. Conf., in Advances in Information & Communication. San Francisco, CA, USA: Springer, 2020, pp. 556–570.

[2] I. Matloob, S. A. Khan, R. Rukaiya, M. A. K. Khattak, & A. Munir, ''A sequence mining-based novel architecture for detecting fraudulent transactions in healthcare systems,'' IEEE Access, vol. 10, pp. 48447–48463, 2022.

[3] H. Feng, ''Ensemble learning in credit card fraud detection using boosting methods,'' in Proc. 2nd Int. Conf. Comput. Data Sci. (CDS), Jan. 2021, pp. 7–11.

[4] M. S. Delgosha, N. Hajiheydari, & S. M. Fahimi, ''Elucidation of big data analytics in banking: A four-stage delphi study,'' J. Enterprise Inf. Manage., vol. 34, no. 6, pp. 1577–1596, Nov. 2021.

[5] M. Puh & L. Brki¢, ''Detecting credit card fraud using selected machine learning algorithms,'' in Proc. 42nd Int. Conv. Inf. Commun. Technol., Electron. Microelectron. (MIPRO), May 2019, pp. 1250–1255.

[6] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, & A. K. Nandi, ''Credit card fraud detection using AdaBoost & majority voting,'' IEEE Access, vol. 6, pp. 14277–14284, 2018.

[7] N. Kumaraswamy, M. K. Markey, T. Ekin, J. C. Barner, & K. Rascati, ''Healthcare fraud data mining methods: A look back & look ahead,'' Perspectives Health Inf. Manag., vol. 19, no. 1, p. 1, 2022.

[8] E. F. Malik, K. W. Khaw, B. Belaton, W. P. Wong, & X. Chew, ''Credit card fraud detection using a new hybrid machine learning architecture,'' Mathematics, vol. 10, no. 9, p. 1480, Apr. 2022.

[9] K. Gupta, K. Singh, G. V. Singh, M. Hassan, G. Himani, & U. Sharma, ''Machine learning based credit card fraud detection—A review,'' in Proc. Int. Conf. Appl. Artif. Intell. Comput. (ICAAIC), 2022, pp. 362–368.

[10] R. Almutairi, A. Godavarthi, A. R. Kotha, & E. Ceesay, ''Analyzing credit card fraud detection based on machine learning models,'' in Proc. IEEE Int. IoT, Electron. Mechatronics Conf. (IEMTRONICS), Jun. 2022, pp. 1–8.

[11] N. S. Halvaiee & M. K. Akbari, ''A novel model for credit card fraud detection using artificial immune systems,'' Appl. Soft Comput., vol. 24, pp. 40–49, Nov. 2014.

[12] A. C. Bahnsen, D. Aouada, A. Stojanovic, & B. Ottersten, ''Feature engineering strategies for credit card fraud detection,'' Expert Syst. Appl., vol. 51, pp. 134–142, Jun. 2016.

[13] U. Porwal & S. Mukund, ''Credit card fraud detection in e-commerce: An outlier detection approach,'' 2018, arXiv:1811.02196.

[14] H. Wang, P. Zhu, X. Zou, & S. Qin, ''An ensemble learning framework for credit card fraud detection based on training set partitioning & clustering,'' in Proc. IEEE SmartWorld, Ubiquitous Intell. Comput., Adv. Trusted Comput., Scalable Comput. Commun., Cloud Big Data Comput., Internet People Smart City Innov. (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Oct. 2018, pp. 94–98.

[15] F. Itoo, M. Meenakshi, & S. Singh, ''Comparison & analysis of logistic regression, Naïve Bayes & knn machine learning algorithms for credit card fraud detection,'' Int. J. Inf. Technol., vol. 13, no. 4, pp. 1503–1511, 2021.

[16] T. A. Olowookere & O. S. Adewale, ''A framework for detecting credit card fraud among cost-sensitive meta-learning ensemble approach,'' Sci. Afr., vol. 8, Jul. 2020, Art. no. e00464.

[17] A. A. Taha & S. J. Malebary, ''An intelligent approach towards credit card fraud detection using an optimized light

gradient boosting machine,'' IEEE Access, vol. 8, pp. 25579–25587, 2020.

[18] X. Kewei, B. Peng, Y. Jiang, & T. Lu, ''A hybrid deep learning model for online fraud detection,'' in Proc. IEEE Int. Conf. Consum. Electron. Comput. Eng. (ICCECE), Jan. 2021, pp. 431–434.

[19] T. Vairam, S. Sarathambekai, S. Bhavadharani, A. K. Dharshini, N. N. Sri, & T. Sen, ''Evaluation of Naïve Bayes & voting classifier algorithm for credit card fraud detection,'' in Proc. 8th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS), Mar. 2022, pp. 602–608.

[20] P. Verma & P. Tyagi, ''Analysis of supervised machine learning algorithms in context of fraud detection,'' ECS Trans., vol. 107, no. 1, p. 7189, 2022.

[21] J. Zou, J. Zhang, & P. Jiang, ''Credit card fraud detection using autoencoder neural network,'' 2019, arXiv:1908.11553.

[22] D. Almhaithawi, A. Jafar, & M. Aljnidi, ''Example-dependent costsensitive credit cards fraud detection using SMOTE & Bayes minimum risk,'' Social Netw. Appl. Sci., vol. 2, no. 9, pp. 1–12, Sep. 2020.

[23] J. Cui, C. Yan, & C. Wang, ''Learning transaction cohesiveness for online payment fraud detection,'' in Proc. 2nd Int. Conf. Comput. Data Sci., Jan. 2021, pp. 1–5.

[24] M. Rakhshaninejad, M. Fathian, B. Amiri, & N. Yazdanjue, ''An ensemble-based credit card fraud detection algorithm using an efficient voting strategy,'' Comput. J., vol. 65, no. 8, pp. 1998–2015, Aug. 2022.

[25] A. H. Victoria & G. Maragatham, ''Automatic tuning of hyperparameters using Bayesian optimization,'' Evolving Syst., vol. 12, no. 1, pp. 217–223, Mar. 2021.

[26] H. Cho, Y. Kim, E. Lee, D. Choi, Y. Lee, & W. Rhee, ''Basic enhancement strategies when using Bayesian optimization for hyperparameter tuning of deep neural networks,'' IEEE Access, vol. 8, pp. 52588–52608, 2020.

[27] F. N. Khan, A. H. Khan, & L. Israt, ''Credit card fraud prediction & classification using deep neural network & ensemble learning,'' in Proc. IEEE Region 10 Symp. (TENSYMP), Jun. 2020, pp. 114–119.

[28] W. Liang, S. Luo, G. Zhao, & H. Wu, ‘‘Predicting hard rock pillar stability using GBDT, XGBoost, & LightGBM algorithms,’’ Mathematics, vol. 8, no. 5, p. 765, May 2020.

[29] S. B. Jabeur, C. Gharib, S. Mefteh-Wali, & W. B. Arfi, ‘‘CatBoost model & artificial intelligence techniques for corporate failure prediction,’’ Technol. Forecasting Social Change, vol. 166, May 2021, Art. no. 120658.

[30] J. Hancock & T. M. Khoshgoftaar, ‘‘Medicare fraud detection using CatBoost,’’ in Proc. IEEE 21st Int. Conf. Inf. Reuse Integr. Data Sci. (IRI), Aug. 2020, pp. 97–103.

[31] B. Dhananjay & J. Sivaraman, ‘‘Analysis & classification of heart rate using CatBoost feature ranking model,’’ Biomed. Signal Process. Control, vol. 68, Jul. 2021, Art. no. 102610.

[32] Y. Chen & X. Han, ‘‘CatBoost for fraud detection in financial transactions,’’ in Proc. IEEE Int. Conf. Consum. Electron. Comput. Eng. (ICCECE), Jan. 2021, pp. 176–179.

[33] A. Goyal & J. Khiari, ‘‘Diversity-aware weighted majority vote classifier for imbalanced data,’’ in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Jul. 2020, pp. 1–8.

[34] A. Roy, J. Sun, R. Mahoney, L. Alonzi, S. Adams, & P. Beling, ‘‘Deep learning detecting fraud in credit card transactions,’’ in Proc. Syst. Inf. Eng. Design Symp. (SIEDS), Apr. 2018, pp. 129–134.