iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.42829

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Detection of Fake Currency Using KNN Algorithm

Dr. P V Kumar[1], V. Akhila[2], B. Sushmitha[3], K. Anusha[4]
[1, 2]*Professor, Dept of IT, Anurag Group of Institutions, Telangana*
[1, 2, 3, 4]*Bachelor of Technology, Information Technology, Anurag Group of Institutions, Telangana, India*

Abstract: ***Bank currency is our country's most valuable asset, and to cause inconsistencies in money, criminals use phony notes that seem identical to the real one on the stock exchange. During demonetization time it is seen that so much fake currency is floating in the market. In general, for a human being, it is very difficult to identify forged notes from the genuine not instead of various parameters designed for identification as many features of forged notes are similar to the original one. To discriminate between fake bank currency and original note is a challenging task. So, there must be an automated system that will be available in banks or ATMs. To design such an automated system there is a need to design an efficient algorithm that can predict whether the banknote is genuine or forged bank currency as fake notes are designed with high precision. In this paper, six supervised machine learning algorithms are applied to the dataset available on the UCI machine learning repository for the detection of Bank currency authentication. To implement this we have applied Support Vector Machine, Random Forest, Logistic Regression, Naïve Bayes, Decision Tree, K- Nearest Neighbor by considering three train test ratios 80:20, 70:30, and 60:40 and measured their performance based on various quantitative analysis parameters like Precision, Accuracy, Recall, MCC, F1-Score and others. And some SML algorithms are giving 100 % accuracy for a particular train test ratio.***

Keywords: Support Vector Machine, Bank currency, Supervised Machine Learning

I. INTRODUCTION

Financial activities are carried out every second by many persons in which one most important asset of our country is Banknotes [3]. Fake notes are introduced in the market to create discrepancies in the financial market, even if they resemble the original note. They are illegally created to complete the various task [12]. In 1990 forgery issue is not much of a concern but in the late 19th-century forgery has been increasing drastically [13]. In the 20th century, technology is increasing very vastly that will help the frauds to generate fake notes whose resemblance is genuine not and it is very difficult to discriminate against them [1]. This will lead the financial market to its lowest level. To stop this and to conduct smooth transaction circulation forged bank currency must be conserved [16]. As a human being, it is very difficult to identify between genuine and forged bank currency. The government has designed banknotes with some features by which we can identify genuine [9]. But frauds are creating fake notes with almost the same features with nice accuracy that make it very difficult to identify genuine notes [5]. So, nowadays it is required that banks or ATMs must have some system that can identify the forged note from the genuine note [12]. To determine the legitimacy of the banknote artificial intelligence and Machine learning(ML) can play a vital role to design such a system that can identify forged notes from the genuine bank currency[6,7,12]. Nowadays, supervised machine learning (SML) approaches for classification problem is widely used. For medical diseases its shows even promising results [2]. Few authors have only applied SML algorithms to bank currency authentication [6-9, 12]. To identify whether a note is genuine or fake we have to develop an automation system. Initially, the input is an image of a note and from different image processing techniques, we can extract the features of the note. Further, these features are given as input to the SML algorithms to predict whether a note is original or fake. In review, we can see that not much work is done on this site. Contribution of the paper: First we have visualized the dataset taken from the UCI ML repository using different types of plotting, and pre-processed the data. Further, SML algorithms Decision tree (DT), KNN, and SVM are applied to the data set which contains the features extracted from the bank currency to classify them as to whether it is original or not. For analysis of their result, we have applied SML algorithms on the dataset with three different train test ratio and their results are compared based on different SML algorithms standard evaluation parameters like MCC, F1 Score, NPV, NDR, accuracy, and others. In section II literature review of papers is discussed followed by that in Section III description of the dataset is given. Further Results of various classification algorithms are analyzed based on standard quantitative analysis parameters and Conclusions are drawn in Section IV.

©IJRASET: All Rights are Reserved

II. LITERATURE SURVEY

In this section, a review of some papers is discussed those applied machine learning approaches to classify whether a note is original or not. Yeh et. al. implemented SVM based on multiple kernels to reduce the false rate and compare d with SVM (single kernel) [17,19]. To classify real and forged networks. Author Hassanpour et. al. used a texture-based feature extraction method for the recognition and to model texture Markov chain concept is used. This method can recognize different countries' currencies [5]. To classify whether the note is forged or not global optimization algorithms are applied in Artificial Neural Network (ANN) training phase, and they have observed success in the classification of notes [8, 14, 15, 11]. Decision tree and MLP algorithms are used to classify the bank currency in [7]. Further multi-classification was done using wavelet for feature extraction by [4] BPN and SVM machine learning algorithms are used to classify the bank currency and it's found that BPN is giving more accuracy than SVM. [6, 16]. In [2] for the counterfeit type of currency notes classification is done using segmentation for the feature extraction based on different regions of the note. The same type of study is done in [6] where bank currency features are extracted using segmentation of image and further these features are given as input to SVM for determining the note authentication. A neural network (NN) is applied to the Thai bank currency for classification, a scanner is used to collect the note image and to covert in the bit map for feature extraction, and then these data are given to BPN for detecting authentication of Bank currency [15]. The probabilistic NN method is used for the classification of bank currency [9] and in [10] LVQ classifier is used for detecting note authentication. Both the authors of the paper applied the above approaches to the US Dollars data set. Recognition of euro banknotes has been proposed by using perceptron of three-layer and to classify bank currency into a particular class by considering input as an image of bank currency The model is trained using the backpropagation method... Further for validation radial basis function is used to discard the invalid data [11].

III. DESCRIPTION OF DATA SET AND RESULTS OF A MACHINE LEARNING ALGORITHM FOR PREDICTION OF FORGED AND GENUINE BANK CURRENCY

Description of the dataset: Dataset is taken from the machine learning repository of UCI to train the models [18]. The features of data are extracted from the forge and genuine images of banknotes. The total instance in the dataset are 1372 and 5 attributes are present. In dataset 4 are the features and one is the target. The dataset is divided into two classes forge and genuine, ratio (55: 45 percent) of both the classes is balanced. Two values are present in the target attribute i.e., 0 and 1 where 1 represents the fake note and 0 is represented a genuine note. In this section, the results of various SML algorithm is discussed in detail. SVM, KNN, and DT are applied to bank currency authentication data to classify whether the note is genuine or forged. To accomplish this task three train test ratios are considered 80:20, 60:40, and 70:30, and further above algorithms are applied to test their accuracy and to also derive various other quantitative analysis parameters for evaluation of the ML models' performance. The following results are observed after applying various SML algorithms:

An SML algorithms description with the ROC and Learning curves to measure accuracy for train test ratio 80:20

SVM: It is an SML model to classify the data based on pattern recognition. To separate the two classes a decision boundary is created in the data. Dataset items are plotted on the graph then classification is performed for differentiating the two classes using the hyperplane concept. The kernel function is used to convert linearly separable data from the non-linearly separable data. For less number of features linear kernel is used for classification and number of large test cases [6]. SVM is applied to the dataset by considering three different train test ratios (80:20, 60:40, and 70:30) to predict whether the bank currency is forged or genuine. For the train test ratio, 80:20 ROC curve and learning curves are drawn. From Fig. 1. The accuracy of SVM is observed at around 98% see Fig. 1.

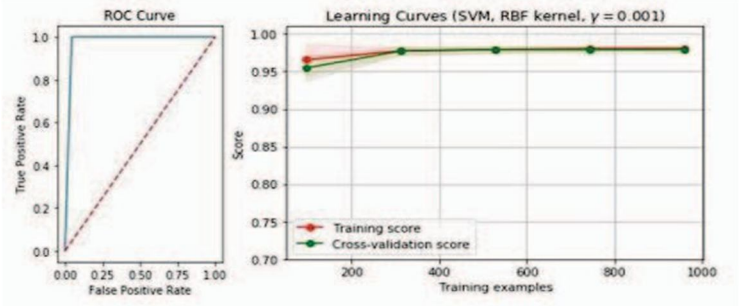


Fig. 1. SVM ROC and Learning curve for training test ratio 80:20

DT: It is a classification model having a structure like a tree. DT is incrementally developed by breaking down the data set into smaller subsets. DT results is having two types of nodes Decision nodes and leaf nodes. For an example consider a decision node i.e., Outlook and it have branches such as Rainy, Overcast and Sunny representing values of the tested feature. Hours Played i.e., a leaf node it gives the decision on numerical targeted value. DT can handle both numerical as well as categorical data [8]. DT is applied to the dataset by considering three different train test ratios (80:20, 60:40, and 70:30) to predict whether the bank currency is forged or genuine. For the train test ratio, 80:20 ROC curve and learning curves are drawn. See Fig. 2. The accuracy of DT has been observed at around 99% see Fig. 2.

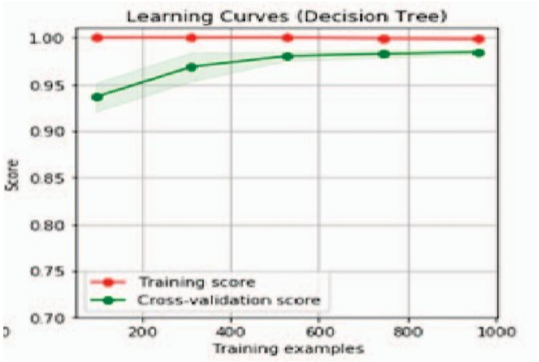


Fig. 2. DT ROC and Learning curve for train test ratio 80:20.

KNN: It is an SML model that may be used for classification as well as regression problems of prediction but mainly in the industry it is used for classification problems. KNN is a lazy algorithm means it learns very slowly as its training is very slow due to the consideration of the whole dataset for classification. And it is also known as a parametric learning algorithm as it will not consider any information from the underlying data. KNN uses the concept of feature similarity to find out the new data point values [9] i.e., the value assigned to the new data point is based on the matching of its value to the points of the training set [9]. KNN is applied to the dataset by considering three different train test ratios (80:20, 60:40, and 70:30) to predict whether the bank currency is forged or genuine. For the train test ratio, an 80:20 ROC curve and learning curves are drawn see Fig. 3. The accuracy of KNN is observed at around 100% see Fig. 3.

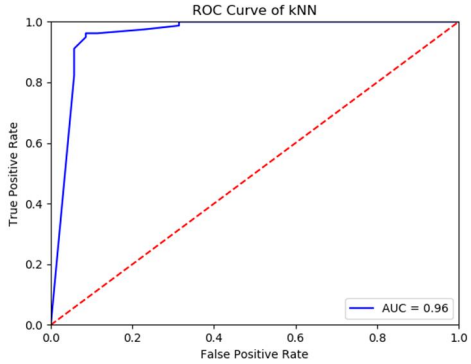


Fig. 3. ROC Curve for KNN

IV. EXISTING SYSTEM

Nowadays, supervised machine learning (SML) approaches for classification problem is widely used. For medical diseases its shows even promising results [2]. Few authors have only applied SML algorithms to bank currency authentication [6-9, 12]. To identify whether a note is genuine or fake we have to develop an automation system. Initially, the input is an image of a note and from different image processing techniques, we can extract the features of the note. Further, these images are given as input to the SML algorithms to predict whether the note is original or fake. In review, we can see that not much work is done on this site**.**

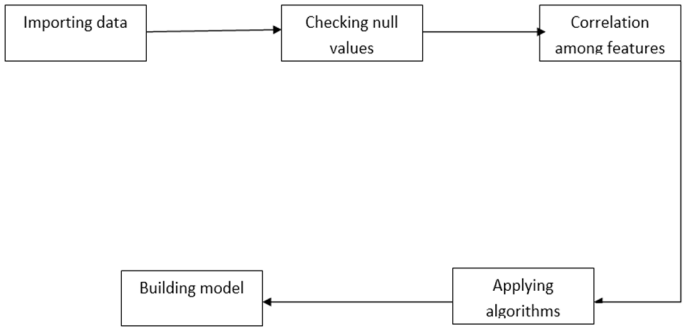
A few of the demerits of the existing system are it is very difficult to identify forged notes from the genuine and it is difficult to discriminate between fake bank currency and original note is a challenging task.

V. PROPOSED SYSTEM

To discriminate between fake bank currency and original note is a challenging task. So, there must be an automated system that will be available in banks or ATMs. To design such an automated system there is a need to design an efficient algorithm that can predict whether the banknote is genuine or forged bank currency as fake notes are designed with high precision. In this project, three supervised machine learning algorithms are applied to the dataset available on the UCI machine learning repository for the detection of Bank currency authentication. To implement this we have applied the Support Vector Machine, Decision Tree, K- Nearest Neighbor by considering three train test ratios 80:20, 70:30, and 60:40. Some of the advantages of the proposed system are Effectiveness and High performance.

VI. ARCHITECTURE

The architecture provides the entire process flow of the system.



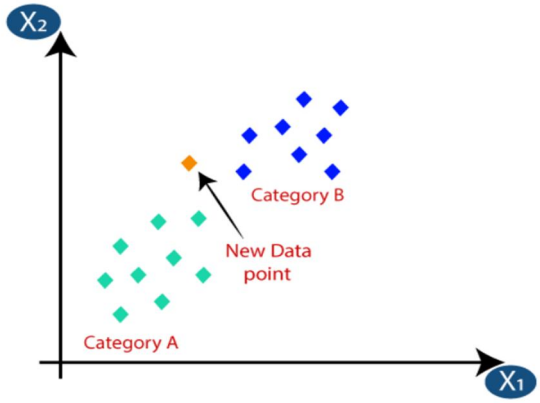
VII. IMPLEMENTATION

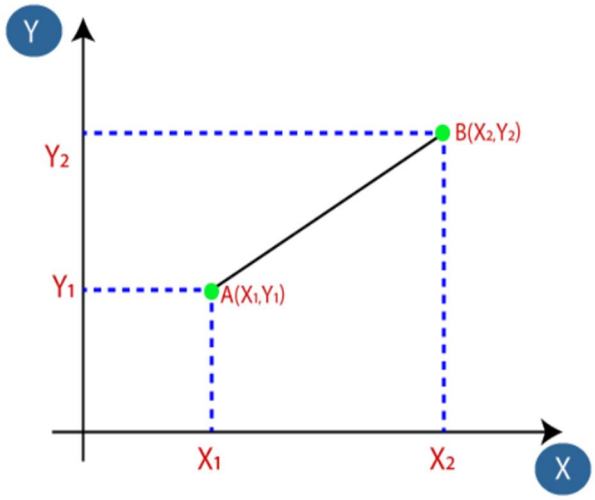
A. *Algorithms Used*

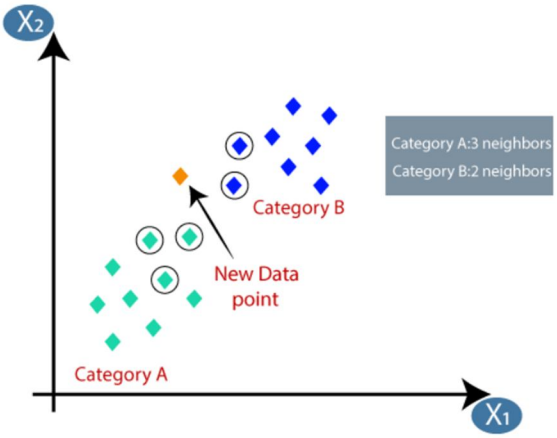
First, we have visualized the dataset taken from the UCI ML repository using different types of plotting, and pre-processed the data. Further, SML algorithms Decision tree (DT), KNN, and SVM are applied to the data set which contains the features extracted from the bank currency to classify them as to whether it is original or not. For analysis of their result, we have applied SML algorithms to the dataset with three different train test ratios.

1) KNN Algorithm

The k-nearest neighbors (KNN) algorithm is a simple, easy-to-implement supervised machine learning algorithm that can be used to solve both classification and regression problems. Machine learning models use a set of input values to predict output values. KNN is one of the simplest forms of machine learning algorithms mostly used for classification. It classifies the data point on how its neighbor is classified. KNN classifies the new data points based on the similarity measure of the earlier stored data points. For example, if we have a dataset of tomatoes and bananas. KNN will store similar measures like shape and color. When a new object comes it will check its similarity with the color (red or yellow) and shape.

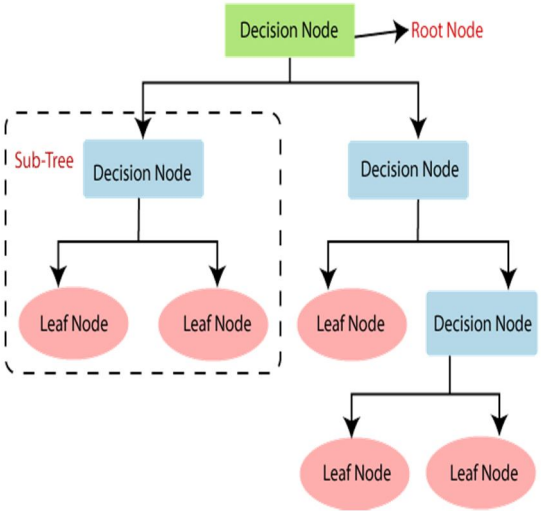






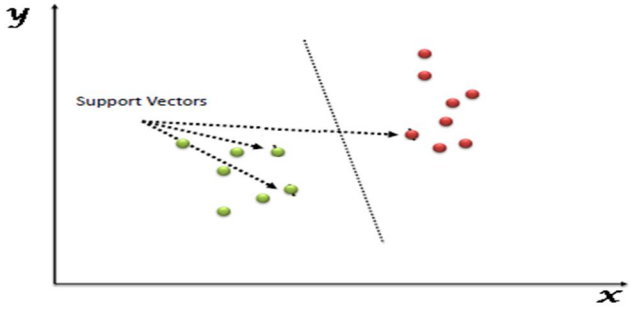
2) Decision Tree Algorithm

The Decision Tree algorithm belongs to the family of supervised learning algorithms. Unlike other supervised learning algorithms, the decision tree algorithm can be used for solving regression and classification problems too. The goal of using a Decision Tree is to create a training model that can use to predict the class or value of the target variable by learning simple decision rules inferred from prior data(training data). In Decision Trees, for predicting a class label for a record we start from the root of the tree. We compare the values of the root attribute with the record's attribute.



3) Support Vector Machine

SVM stands for supervised machine learning and can be used to solve classification or regression problems. It transforms your data using a technique known as the kernel trick and then calculates an ideal boundary between the possible outputs based on these alterations. The kernel trick alters the data you provide it with. Outcomes some data that you don't recognize, and in goes some excellent traits that you hope will form a terrific classifier. It's similar to unraveling a DNA strand. You start with a seemingly innocent vector of data, which is unraveled and compounded into a considerably larger set of data after going through the kernel trick.

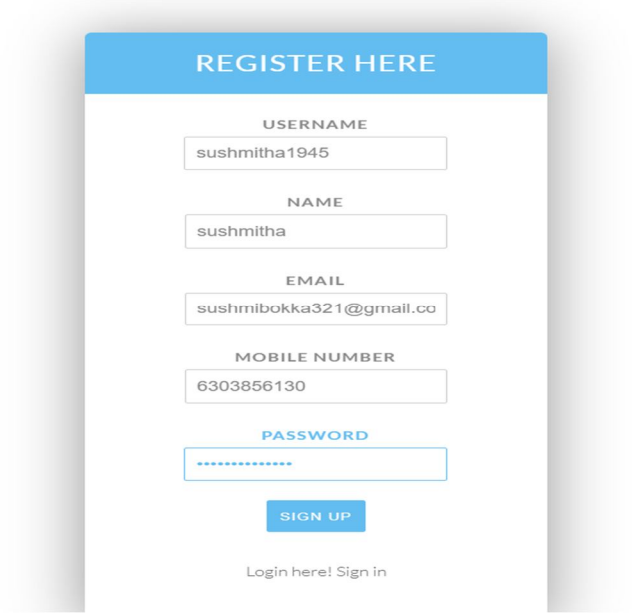


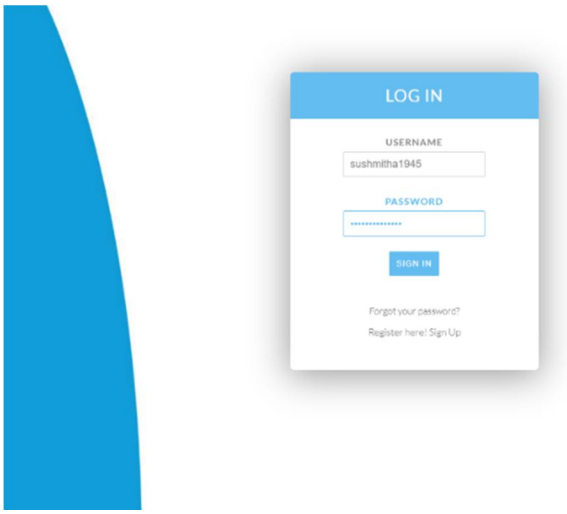
VIII. RESULTS

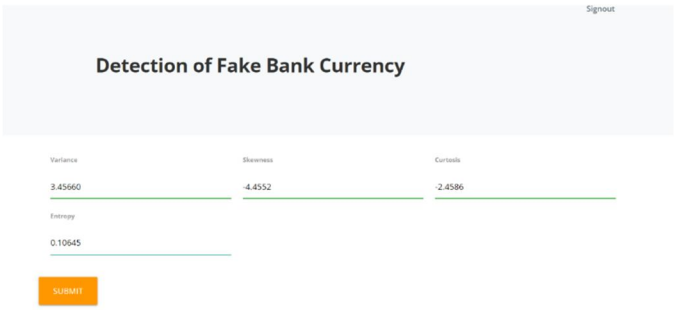
The proposed system was successfully tested to denote its effectiveness and achievability. Thus a fake note can be easily differentiated from a genuine note.

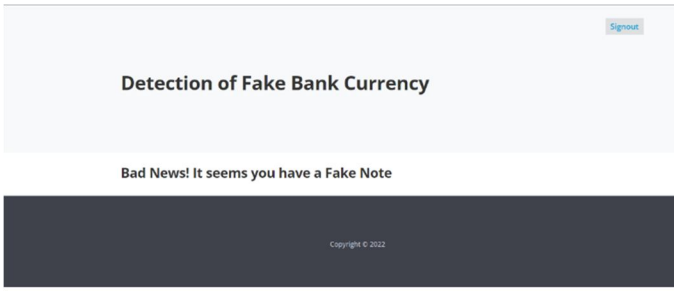


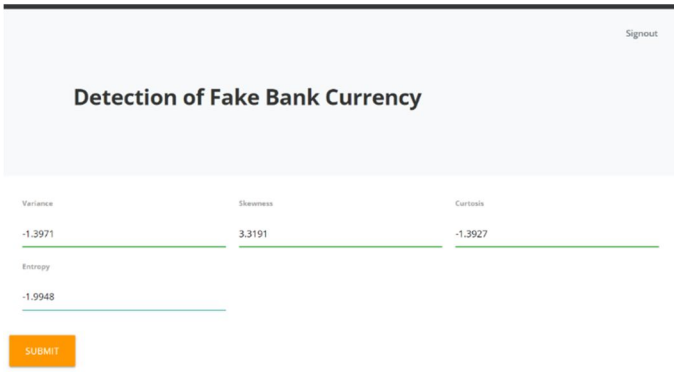


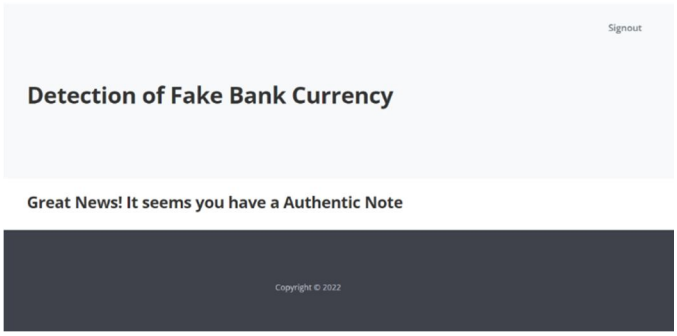



LOG IN
USERNAME
sushmitha1945
PASSWORD
SIGN IN
Forgot your password?
Register here! Sign Up


Signout
Detection of Fake Bank Currency
Variance
3.45660
Skewness
-4.4552
Curtosis
-2.4586
Entropy
0.10645
SUBMIT


Signout
Detection of Fake Bank Currency
Bad News! It seems you have a Fake Note
Copyright © 2022


Signout
Detection of Fake Bank Currency
Variance
-1.3971
Skewness
3.3191
Curtosis
-1.3927
Entropy
-1.9948
SUBMIT



IX. CONCLUSION

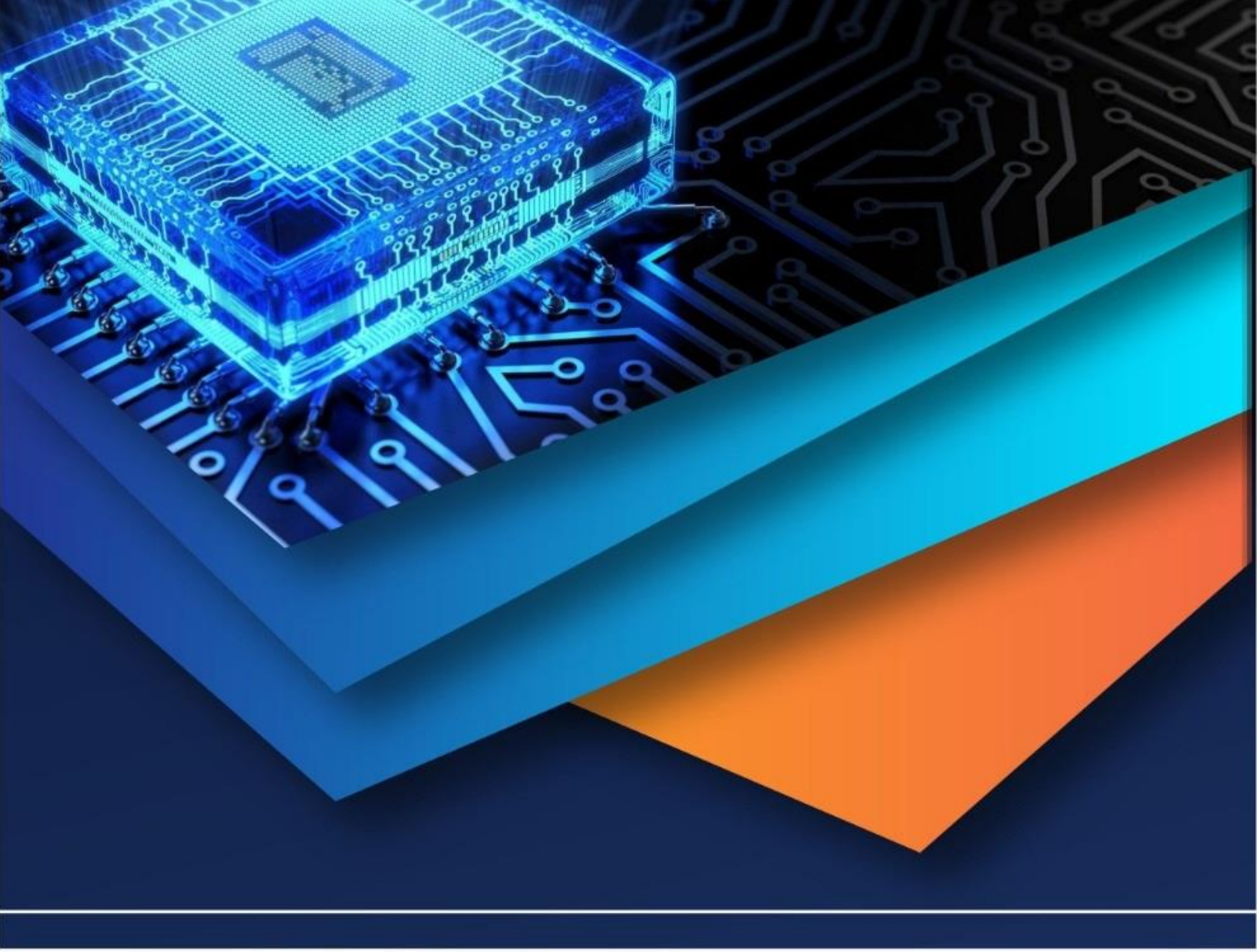
In this project, SML algorithm SVM, DT, and KNN are applied to the banknote authentication dataset taken from the UCI ML repository on three different train test ratios 80:20, 60:40, and 70:30. The dataset contains 1372 and 5 attributes and out of which 4 are the features and one is the target attribute that has value as genuine bank currency or forged note.

REFERENCES

[1] M. Aoba, T. Kikuchi, and Y. Takefuji, "Euro Banknote Recognition System Using a Three-layered Perceptron and RBF Networks", IPSJ Transactions on Mathematical Modeling and it's Applications, May 2003.
[2] S. Desai, S. Kabade, A. Bakshi, A. Gunjal, M. Yeole, "Implementation of Multiple Kernel Support Vector Machine for Automatic Recognition and Classification of Counterfeit Notes", International Journal of Scientific & Engineering Research, October-2014.
[3] C. Gigliaranoa, S. Figini, P. Muliere, "Making classifier performance comparisons when ROC curves intersect", Computational Statistics and Data Analysis 77 (2014) 300–312.
[4] E. Gillich and V. Lohweg, "Banknote Authentication", 2014.
[5] H. Hassanpour and E. Hallajian, "Using Hidden Markov Models for Feature Extraction in Paper Currency Recognition.
[6] Z. Huang, H. Chen, C. J. Hsu, W. H. Chen and S. Wuc, "Credit rating analysis with support vector machines and neural network: a market comparative study", 2004.
[7] C. Kumar and A. K. Dudyala, "Banknote Authentication using Decision Tree rules and Machine Learning Techniques", International Conference on Advances in Computer Engineering and Applications(ICACEA), 2015.
[8] M. Lee and T. Chang, "Comparison of Support Vector Machine and Back Propagation Neural Network in Evaluating the Enterprise Financial Distress", International Journal of Artificial Intelligence & Applications 1.3 (2010) 31-43.
[9] C. Nastoulis, A. Leros, and N. Bardis, "Banknote Recognition Based On Probabilistic Neural Network Models", Proceedings of the 10th WSEAS International Conference on SYSTEMS, Vouliagmeni, Athens, Greece, July 10-12, 2006.
[10] S. Omatu, M. Yoshioka and Y. Kosaka, "Bankcurrency Classification Using Neural Networks", IEEE, 2007.
[11] A. Patle and D. S. Chouhan, "SVM Kernel Functions for Classification", ICATE 2013.
[12] E. L. Prime and D. H. Solomon, "Australia's plastic banknotes: fighting counterfeit currency.," Angewandte Chemie (International ed. in English), vol. 49, no. 22, pp. 3726–36, May 2010.
[13] A. Roy, B. Halder, and U. Garain, "Authentication of currency notes through printing technique verification," Proceedings of the Seventh Indian Conference on Computer Vision, Graphics and Image Processing -ICVGIP '10, pp. 383–390, 2010.
[14] P. D. Shahare and R. N. Giri, "Comparative Analysis of Artificial Neural Network and Support Vector Machine Classification for Breast Cancer Detection", International Research Journal of Engineering and Technology, Dec-2015.
[15] F. Takeda, L. Sakoobunthu and H. Satou, "Thai Banknote Recognition Using Neural Network and Continues Learning by DSP Unit", International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, 2003.

ACKNOWLEDGMENT

We express our sincere gratitude to our guide, Professor **Dr. P V Kumar** for his suggestions and support during every stage of this work. We also convey our deep sense of gratitude to Professor **Dr. K. S. Reddy**, Head of the Information Technology department.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)