FAKE CURRENCY DETECTION USING NOVEL RANDOM FOREST ALGORITHM AND DECISION TREE CLASSIFIER WITH IMPROVED ACCURACY

Renangi Supriya[1], S.Kalaiarasi[2*]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: To enhance the accuracy in classifying fake currency detection using Novel Random Forest Algorithm and Decision Tree Classifier. Materials and Methods: This study contains 2 groups such as Novel Random Forest Algorithm and Decision Tree Classifier. Each group consists of a sample size of 10 and the study parameters include alpha value 0.05, beta value 0.2. SPSS was used for predicting significance value of the dataset considering G-Power value 80%. Their accuracies are compared with each other using different sample sizes. Results and Discussions: The Novel Random Forest Algorithm with accuracy 80.5%, is more accurate than the Decision Tree Classifier with accuracy value 47.5% in classifying the fake currency notes with significance value 0.001 ($p<0.05$). Conclusion: The Random Forest Algorithm is significantly better than the Decision tree classifier in identifying fake notes. It can be also considered as a better option for the classification of fake currency.

Keywords: Fake Currency Detection, Novel Random Forest Algorithm, Decision Tree Classifier, Machine learning, Regression, Image processing.

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and technical Science, Saveetha University, Chennai, Tamil Nadu, India. Pincode: 602105.

[2*]Project Guide, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India. Pincode: 602105.

1. Introduction

In the banking sector, the greatest risk is the generation of fake coins. Most of the time, UltraViolet light is used to prove authenticity. The main features for detecting fake coins are the banknote value, ink stain, security thread, serial number, intaglio printing, watermark, reserve bank number field, marking, topography, micro-writing, as well as numbers and alignment. These key features include watermark, ink blot, security thread, topography, numbers and location, and micro-writing. (National Research Council et al. 2002) However, the following steps generally need to be performed by researchers for machine evaluation. Recursive feature elimination is used to choose either the best or worst performing feature and a repeatedly pruned set is performed. Fake currency perceived judgment can be made precise and effective using the algorithm. This recursive feature elimination uses a part of linear regression. The correct extraction feature is selected and trained with the Random Forest algorithm (RFA). (Radbruch 2013) Transformation of the original input set to higher dimensional feature space is done using the kernel function of RFA (Bhatia et al. 2021), in order to use the hyperlink which is required for the RFA algorithm. The advantage of RFA is that it requires less training data when compared to other models. The detection of currency notes data sets improves the accuracy by a greater rate. The proposed algorithm helps in improving the accuracy of fake notes. (National Research Council et al. 2002; Bartram and Ballance 2020). The fake currency detection is applicable in the Reserve Bank of India and other banking sectors. Fake currency detection is a serious problem that affects almost every country's economy, including India's. Currency duplication, also known as counterfeit currency (United States. Secret Service 1975), is a serious economic threat. Because of advances in printing and scanning technology, it is now a common occurrence.

There are about 34 articles in IEEE Xplore and 10 in Scopus related to this study. RFA is derived from the usage of sequential data information (Shu and Liu 2019). The detection of currency and image recognition based on deep learning was performed. Analysis of fake currency detection techniques for classification models were developed. The Report was prepared on fake currency detection using Image Processing Method. Automatic Cash Deposit Machine was developed and Currency Detection was implemented with the help of Fluorescent and UV Light (Bátiz-Lazo 2018). The Fake currency detection using feature extraction was reviewed elaborately (Bleay, Croxton, and De Puit 2018).Our team has extensive knowledge and research experience that has translated into high quality publications(Pandiyan et al. 2022; Yaashikaa, Devi, and Kumar 2022; Venu et al. 2022; Kumar et al. 2022; Nagaraju et al. 2022; Karpagam et al. 2022; Baraneedharan et al. 2022; Whangchai et al. 2022; Nagarajan et al. 2022; Deena et al. 2022)

Some datasets are intended for theoretical research rather than processing for real-world applications. The drawback is that defining the boundaries between the fake notes and original notes is extremely difficult. Because most existing standard feature extraction processes are designed for short-term analysis, researchers created their own feature set. (U. s. Department of Justice 2014). The aim is to improve accuracy in detection of the fake currencies.

2. Materials and Methods

The number of required samples in research are two in which group 1 is RFA compared with group 2 of DTC Classifier. The samples were taken from the device and iterated 10 times to get desired accuracy with G-power 80%, threshold 0.05% and CI 95%. A dataset consisting of a collection of booknotes was downloaded from the Data sci-hub repository (Datopian 2014) and the comparative study of the two groups of algorithms is made for improving accuracy in fake currency detection.

The data collection is taken from the open source access website IEEE-dataport.org that is used for software effort estimation using Novel Random Forest and Decision Tree technique. The open access dataset consists of 108 rows and 10 columns. The Jupyter software with windows 10.1 system has been used to develop this software effort estimation. The proposed system uses two groups: the Novel Random Forest and Decision Tree technique where these algorithms are fitted into the dataset which is then tested and trained for the process of estimating the software effort where the cost estimation and the time estimation is known.

Novel Random Forest Algorithm

A Novel Random Forest Algorithm (RFA) is often used in the fake currency to predict future profits. It has a big effect on economic system forecasting. So, the program predicts the fake notes.

Pseudocode for Novel Random Forest Algorithm

Input: Training dataset K.

Output: A class of testing dataset

Step 1: Choose K data points at random from the training set.
Step 2: Create decision trees for the data points you've chosen (Subsets).
Step 3: Choose a N for the number of decision trees you want to make.
Step 4: Repeat of Steps 1 and 2.
Step 5: Find the forecasts of each decision tree for new data points, and allocate the new
data points to the category with the most votes.

Decision Tree Classifier
Decision Tree is a supervised learning strategy for tackling classification and regression problems, but it is most typically employed for classification problems (Irizarry 2019). Internal nodes contain dataset attributes, branches represent decision rules, and each leaf node provides the result in this tree-structured classifier that incorporates image processing.

Pseudocode for Decision Tree Classifier
Input: K is the training dataset.
Output:A class of testing dataset
Step 1: Start with the root node, which holds the entire dataset, explains S.
Step 2: Using the Attribute Selection Measure, find the best attribute in the dataset (ASM).
Step 3: Subdivide the S into subsets that contain the best attribute's possible values.
Step 4: Create the node of the decision tree that has the best attribute.
Step 5: Create additional decision trees in a recursive manner using the subsets of the dataset obtained in step
3. Continue this process until the nodes can no longer be classified, at which point the final node is referred to as a leaf node.

Recall that the testing setup includes both hardware and software configuration choices. The laptop has an Intel Core i3 7th generation CPU with 12GB of RAM, an x86-based processor, a 64-bit operating system, and a hard drive. Currently, the software runs on Windows 10 and is programmed in Python. Once the program is finished, the accuracy value will appear. Procedure: Wi-Fi laptop connected. Chrome to Google Collaboratory search Write the code in Python. Run the code. To save the file, upload it to the disc, and create a folder for it. Log in using the ID from the message. Run the code to output the accuracy and graph.

Statistical Analysis
The proposed system utilized 10 iterations for each group with predicted accuracy noted and analyzed. Independent samples t-test was done to obtain significance between two groups. In fake currency detection parameters are independent variables and fake prediction is dependent variable.

3. Results

Table 1 shows the accuracy value of iteration of RFA and DTC. Table 2 represents the Group statistics results which depicts RFA with mean accuracy of 65.50%, and standard deviation is 3.028. DTC has a mean accuracy of 47.50% and standard deviation is 3.028. Proposed RF algorithm provides better performance compared to the LR algorithm. Table 3 shows the independent samples T-test value for RFA and DTC with Mean difference as 8.1, std Error Difference as 0.67. Significance value is observed as 0.001 ($p<0.05$). Fig. 1 shows the bar graph comparison of mean of accuracy on RFA and DTC algorithm. Mean accuracy of RFA is 65% and DTC is 47%.

4. Discussion

In this study, detecting fake currency using the RF algorithm has significantly higher accuracy, approximately 65% in comparison to DTC 47%. RFA appears to produce more consistent results with minimal standard deviation.
The similar findings of the paper had an accuracy of 65% with RFA which was used to detect the currency. The proposed work of reported RFA has 65% accuracy which is used to predict the accuracy of fake currency notes and performance of fake notes. The work proposed by Rathore 2020 shows the RFA has a better accuracy of 65%. DTC is a parameter to measure fake currency which is used in both traditional and modern methods as per their research it opposes DTC has highest accuracy and DTC will get least accuracy compared to other machine learning techniques (Council of Europe 2007) which ranges between 60% when compared to other machine learning algorithms will get more accuracy than this (Harvard Business Review et al. 2019). By using RFA for forecasting fake currency it will have key issues to pretend (Mehta, Pandya, and Kotecha 2021) in this paper shows RFA has the least accuracy of 47%. Increasing the dataset's value only tends to get desired accuracy. RFA performs better with a combination of other machine learning algorithms (Ireton and Posetti 2018).
The limitation of this research is that appropriate results are not obtained for smaller data, and also not all the parameters are considered for training. The future scope of proposed work will be prediction of stock price based on classification using class labels with lesser time complexity.

5. Conclusion

In this research, detection of the fake currency notes was implemented using the RFA and DTC. The accuracy value of the RFA is 65% whereas the accuracy value of DTC is 47%. The quality of detecting fake notes accuracy using RFA appears to be better than DTC.

Declaration
Conflict of Interests
No conflict of interests in this manuscript

Authors Contribution
Author RS was involved in data collection, data analysis, and manuscript writing. Author SKA was involved in conceptualization, data validation, and critical review of manuscript.

Acknowledgement
The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary Infrastructure to carry out this work successfully.

Funding
We thank the following organizations for providing financial support that enabled us to complete the study.

1. HKM Info Systems Pvt.Ltd.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering

6. References

Bartram, Jamie, and Richard Ballance. 2020. Water Quality Monitoring: A Practical Guide to the Design and Implementation of Freshwater Quality Studies and Monitoring Programmes. CRC Press.

Bátiz-Lazo, Bernardo. 2018. Cash and Dash: How ATMs and Computers Changed Banking. Oxford University Press.

Benware, Dewey. 2021. Machine Learning Algorithms: How The Machine Learning Algorithms Work Behind The Scenes: Random Forest Algorithm Geeksforgeeks.

Bhatia, Aman, Vansh Kedia, Anshul Shroff, Mayand Kumar, Bickey Kumar Shah, and Aryan. 2021.

"Fake Currency Detection with Machine Learning Algorithm and Image Processing." In 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE.
https://doi.org/10.1109/iciccs51141.2021.9432274.

Bleay, Stephen M., Ruth S. Croxton, and Marcel De Puit. 2018. Fingerprint Development Techniques: Theory and Application. John Wiley & Sons.

Council of Europe. 2007. Cyberterrorism: The Use of the Internet for Terrorist Purposes. Council of Europe.

Datopian. "Stock Market Data." Accessed August 9, 2021. https://datahub.io/collections/stock-market-data.

Harvard Business Review, Don Tapscott, Marco Iansiti, and Karim R. Lakhani. 2019. Blockchain: The Insights You Need from Harvard Business Review. Harvard Business Press.

Ireton, Cherilyn, and Julie Posetti. 2018. Journalism, Fake News & Disinformation: Handbook for Journalism Education and Training. UNESCO Publishing.

Irizarry, Rafael A. 2019. Introduction to Data Science: Data Analysis and Prediction Algorithms with R. CRC Press.

Mehta, Pooja, Sharnil Pandya, and Ketan Kotecha. 2021. "Harvesting Social Media Sentiment Analysis to Enhance Stock Market Prediction Using Deep Learning." PeerJ. Computer Science 7 (April): e476.

National Research Council, Committee on Earth Resources, Board on Earth Sciences and Resources, National Materials Advisory Board, and Committee on Technologies for the Mining Industries. 2002. Evolutionary and Revolutionary Technologies for Mining. National Academies Press.

Radbruch, Andreas. 2013. Flow Cytometry and Cell Sorting. Springer Science & Business Media.

"Review on Detection of Fake Currency Using Image Processing Techniques.". Accessed September 30, 2021.
https://ieeexplore.ieee.org/document/9432111.

Shu, Kai, and Huan Liu. 2019. Detecting Fake News on Social Media. Morgan & Claypool Publishers.

Thomas Collins, M. 1990. Counterfeit Currency: How to Really Make Money. Loompanics Unltd.

United States. Secret Service. 1975. The Detection of Counterfeit Currency: A Law Enforcement Officer's Guide. U. s.

Baraneedharan, P., Sethumathavan Vadivel, C. A. Anil, S. Beer Mohamed, and Saravanan Rajendran. 2022. "Advances in Preparation, Mechanism and Applications of Various Carbon Materials in Environmental Applications: A Review." Chemosphere. https://doi.org/10.1016/j.chemosphere.2022.134596.

Deena, Santhana Raj, A. S. Vickram, S. Manikandan, R. Subbaiya, N. Karmegam, Balasubramani Ravindran, Soon Woong Chang, and Mukesh Kumar Awasthi. 2022. "Enhanced Biogas Production from Food Waste and Activated Sludge Using Advanced Techniques – A Review." Bioresource Technology. https://doi.org/10.1016/j.biortech.2022.127234.

Karpagam, M., R. Beaulah Jeyavathana, Sathiya Kumar Chinnappan, K. V. Kanimozhi, and M. Sambath. 2022. "A Novel Face Recognition Model for Fighting against Human Trafficking in Surveillance Videos and Rescuing Victims." Soft Computing. https://doi.org/10.1007/s00500-022-06931-1.

Kumar, P. Ganesh, P. Ganesh Kumar, Rajendran Prabakaran, D. Sakthivadivel, P. Somasundaram, V. S. Vigneswaran, and Sung Chul Kim. 2022. "Ultrasonication Time Optimization for Multi-Walled Carbon Nanotube Based Therminol-55 Nanofluid: An Experimental Investigation." Journal of Thermal Analysis and Calorimetry. https://doi.org/10.1007/s10973-022-11298-4.

Nagarajan, Karthik, Arul Rajagopalan, S. Angalaeswari, L. Natrayan, and Wubishet Degife Mammo. 2022. "Combined Economic Emission Dispatch of Microgrid with the Incorporation of Renewable Energy Sources Using Improved Mayfly Optimization Algorithm." Computational Intelligence and Neuroscience 2022 (April): 6461690.

Department of Justice. 2014. The Fingerprint:

Nagaraju, V., B. R. Tapas Bapu, P. Bhuvaneswari, R. Anita, P. G. Kuppusamy, and S. Usha. 2022. "Role of Silicon Carbide Nanoparticle on Electromagnetic Interference Shielding Behavior of Carbon Fibre Epoxy Nanocomposites in 3-18GHz Frequency Bands." Silicon. https://doi.org/10.1007/s12633-022-01825-1.

Pandiyan, P., R. Sitharthan, S. Saravanan, Natarajan Prabaharan, M. Ramji Tiwari, T. Chinnadurai, T. Yuvaraj, and K. R. Devabalaji. 2022. "A Comprehensive Review of the Prospects for Rural Electrification Using Stand-Alone and Hybrid Energy Technologies." Sustainable Energy Technologies and Assessments. https://doi.org/10.1016/j.seta.2022.102155.

Venu, Harish, Ibham Veza, Lokesh Selvam, Prabhu Appavu, V. Dhana Raju, Lingesan Subramani, and Jayashri N. Nair. 2022. "Analysis of Particle Size Diameter (PSD), Mass Fraction Burnt (MFB) and Particulate Number (PN) Emissions in a Diesel Engine Powered by Diesel/biodiesel/n-Amyl Alcohol Blends." Energy. https://doi.org/10.1016/j.energy.2022.123806.

Whangchai, Niwooti, Daovieng Yaibouathong, Pattranan Junluthin, Deepanraj Balakrishnan, Yuwalee Unpaprom, Rameshprabu Ramaraj, and Tipsukhon Pimpimol. 2022. "Effect of Biogas Sludge Meal Supplement in Feed on Growth Performance Molting Period and Production Cost of Giant Freshwater Prawn Culture." Chemosphere 301 (August): 134638.

Yaashikaa, P. R., M. Keerthana Devi, and P. Senthil Kumar. 2022. "Advances in the Application of Immobilized Enzyme for the Remediation of Hazardous Pollutant: A Review." Chemosphere 299 (July): 134390.

Tables and Figures

Table 1.Comparison of accuracy values of Random Forest Algorithm and Decision Tree Classifier in various iterations

S.NO	RFA	DTC
1	65	47
2	64	46
3	63	45

4	62	44
5	61	43
6	60	42
7	59	41
8	58	40
9	57	39
10	56	38

Table 2. Group Statistics Results- RFA has an mean accuracy (65.50%), std.deviation (3.028), whereas for DTC has mean accuracy (47.50%), std.deviation (3.028).

Group Statistics					
	Groups	**N**	**Mean**	**Std deviation**	**Std. Error Mean**
Accuracy	RFA	10	65.50	3.028	.957
	NFA	10	47.50	2.057	.896

Table 3. The Independent sample t-test of the significance level RFA and DTC algorithms results with two tailed significant values (p=0.001).

	Independent Samples Test								
	Levene's Test for Equality of Variances					**T-test for Equality of Means**			
Accuracy	**F**	**Sig**	**t**	**df**	**Sig(2-tailed)**	**Mean Difference**	**Std.Error Difference**	**95% Confidence Interval of the Difference**	
								Lower	**Upper**
Equal variances assumed	0.000	1.000	13.294	18	0.001	18.000	1.354	15.155	20.845
Equal variances not assumed			13.294	18	0.001	18.000	1.354	15.155	20.845

Fig. 1. Bar Graph showing Comparison on mean accuracy of RFA (65%) is better than the DTC (47%). X-axis:(GROUPS) RFA vs DTC algorithm and Y-axis: Mean Accuracy with 2 SD.