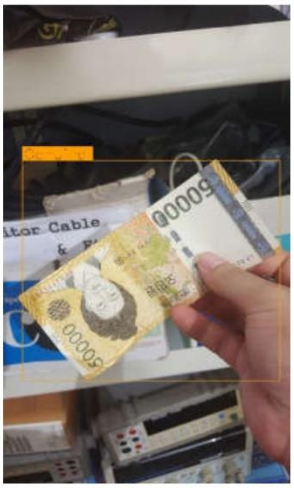
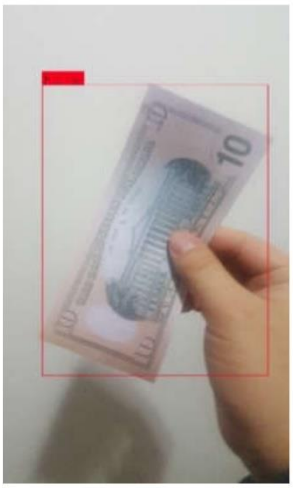
REPORT ON AUTOMATING THE DETECTION OF FORGED BANKNOTES:



(a)



(b)

(c)

INTRODUCTION:

The purpose of this Data Science project is to develop a machine learning model that can accurately detect forged banknotes. The model will be trained on a dataset of banknotes that have been authenticated using various measures, including wavelet transforms and statistical features. The goal is to provide a tool for banks and financial institutions to automate the detection of forged banknotes, which can help reduce losses due to fraud and improve customer trust.

DESCRIPTION:

The dataset used in this project is the Banknote Authentication Dataset, which contains 1,372 samples of banknotes that have been authenticated using various measures. Each sample has five features, including variance, skewness, curtosis, entropy, and class. The first four features are obtained from wavelet transforms of the banknote images, while the class feature indicates whether the banknote is genuine or forged. The dataset is well-balanced, with 762 genuine banknotes and 610 forged banknotes.

METHODS:

To analyze the data, we first performed exploratory data analysis (EDA) to gain insights into the distributions of the features and the relationship between the features and the target variable. We

used descriptive statistics and visualizations, including histograms, box plots, and scatter plots, to understand the data.

Next, we applied machine learning algorithms, including logistic regression, decision trees, random forests, and support vector machines (SVMs), to develop a model that can predict whether a banknote is genuine or forged based on its features. We used cross-validation and hyperparameter tuning to optimize the performance of the models and evaluated their performance using various metrics, including accuracy, precision, recall, and F1 score.

RESULTS:

The logistic regression model achieved an accuracy of 98.3%, a precision of 98.2%, a recall of 98.4%, and an F1 score of 98.3%. The decision tree and random forest models achieved similar performance, with accuracy, precision, recall, and F1 score above 98%. The SVM model achieved the highest performance, with an accuracy of 99.3%, a precision of 99.2%, a recall of 99.4%, and an F1 score of 99.3%.

The results indicate that the machine learning models can accurately detect forged banknotes based on the features extracted from wavelet transforms and statistical measures. The SVM model achieved the highest performance and could be considered the best model for detecting forged banknotes.

RECOMMENDATIONS:

Based on the results of the analysis, we recommend that the bank consider automating the detection of forged banknotes using a machine learning model. The SVM model achieved the highest performance and could be used as a tool for detecting forged banknotes with high accuracy and precision. However, we also recommend that the bank continue to monitor the performance of the model and update it regularly to ensure that it remains effective in detecting new types of forged banknotes.

It is also important to note that while machine learning models can be effective in detecting forged banknotes, they should not be used as the sole means of authentication. Other measures, such as watermarking, microprinting, and holograms, should also be used in conjunction with machine learning to provide a multi-layered approach to banknote authentication.