An Empirical Performance Comparison of Machine Learning Methods for the Detection of Forged Banknotes

Adeel Ahmed and Muhammad Ali

Department of Computer Science, Federal Urdu University of Art, Science and Technology, Karachi
Pakistan

Abstract

This paper seeks to compare the performance of different machine learning methods (stochastic gradient descent, k-nearest neighbors' classifier, decision tree classifier, logistic regression) employed to detect the forgery of banknotes and determine the most suitable method for the detection of forged or counterfeit banknotes.

Keywords: Banknote, KNN Classifier, Decision Tree Classifier, Logistic Regression, Stochastic Gradient Descent, Performance Comparison, Counterfeit Detection

Introduction

Counterfeit currency is one of the major concerns in the financial world. With the advent of the modern technology, it has become a lot easier to forge bank notes that may look as if they are real. As a result, governments around the world have introduced various security features in the bank notes of their respective countries as a preventative measure. But these security features are not completely reliable as they require a keen observer to be aware of them to determine whether a banknote is genuine or forged. Security features present in banknotes are easy to miss by an untrained or uncareful observer. Moreover, visually impaired people cannot take full advantage of these security measures.

In this paper, we seek to detect the genuine and forged banknotes with the help of various machine learning methods, namely Stochastic Gradient Descent, K-Nearest Neighbors' Classifier, Decision Tree Classifier, and Logistic Regression. These algorithms will be applied on the banknote authentication dataset available at [1]. We will then compare the performances of all the employed methods to determine the method most suitable for our dataset.

Literature Review

In [2], Yeh et al. propose the employment of Multiple-Kernal Support Vector Machines for the recognition of counterfeit currency. They utilize the histograms of the captured images as the input for their system.

In [3], Mohamad et al. suggest the use of Artificial Neural Network (ANN) to detect the counterfeit banknotes. Another study [4] by Kaya et al. also suggest the approach of using ANN for the detection of genuine and forged banknotes.

There are also studies which analyze and compare the performance of different machine learning techniques such as [5], in which Shahani et al. employ Back-Propagation Neural Network and Support Vector Machine to compare the performance of these two algorithms for the purpose of banknote authentication.

Similarly, Ghazvini et al. in [6] apply Naïve Bayes and Multilayer Perceptron to the banknote authentication dataset to compare the performances of the aforementioned machine learning methods.

However, to the best of our knowledge, there is no study which applies the machine learning methods: Stochastic Gradient Descent, K-Nearest Neighbors Classifier, Decision Tree, and Logistic Regression to

banknote authentication and compares their performances.

Dataset Description

The banknote authentication dataset can be found at the Machine Learning Repository of University of California, Irvine [1].

The dataset consists of both genuine and forged banknote like specimens. The photos of the both specimens were taken from an industrial camera--- typically used for the print inspection. The result was a greyscale image of 400 x 400 pixels in size with a resolution of about 660 dpi.

Features were extracted from the photos with the application of Wavelet Transform. As shown by Choi et al. in [6], wavelet transform can be effectively used to extract the features from the images of banknotes.

After the extraction of the features, the variance, skewness and kurtosis were calculated of the wavelet-transformed images. This, along with the entropy of the image, make up the four features which determine whether a banknote is genuine or forged. The classification itself is stored as the fifth feature in the dataset ('0' representing a 'genuine' banknote while '1' representing a 'forged' banknote).

The dataset contains 1372 samples: the first 762 being the genuine banknotes whereas the last 610 samples being the forged banknotes.

Experimental Results

To apply and compare the aforementioned algorithms on to the banknote authentication dataset, Python along with its Scikit-learn module [8] was used.

The system with the specifications: Intel Core i5-2400 CPU @ 3.10 GHz & 4 GB RAM was used to conduct the experiment.

For all the experiments, a simple holdout method was used. 200 randomly permuted samples out of the 1372 total samples were reserved for testing purpose. A fixed random seed was used to ensure that the sequence of random permutations remains the same across all the algorithms used for the experiment.

1. Stochastic Gradient Descent

Stochastic Gradient Descent is a simple yet highly efficient approach to discriminative learning of linear classifiers under convex loss functions such as (linear) Support Vector Machines and Logistic Regression.

SGD has been successfully applied to large-scale and sparse machine learning problems often encountered in natural language processing and classification of text.

The efficiency of Stochastic Gradient Descent as well as its ease of implementation (which enables the option for parameter tuning in code) are two of the biggest advantages of this machine learning technique. On the other hand, stochastic gradient descent also has a few disadvantages of Stochastic Gradient Descent which include (but are not limited to): requiring several hyperparameters such as the number of iterations, algorithm itself being sensitive to feature scaling, and the regularization parameter.

After applying the Stochastic Gradient Descent algorithm, following results were obtained:

Accuracy Score	0.99
Time taken by the algorithm to complete*	0.0027 seconds

*Note: Time taken may vary by 0.0001 – 0.0004 seconds.

2. K-Nearest Neighbors Classifier

Neighbors-based classification is a type of *instance-based learning* or *non-generalizing learning* as it does not assume a general model internally, but merely stores the instances of training data. The nearest neighbors of all the individual points contribute to a simple majority vote which aids with the computation of classification. The value k defines the number of nearest neighbors that must be taken into account.

The optimal choice of the value k is highly data-dependent: typically, the effect of noise is suppressed by using a higher value of k, but higher value of k also causes the classification boundaries to be less distinct.

In binary (two class) classification problems, it is helpful to choose *k* to be an odd number as this avoids tied votes.

For the banknote authentication dataset, the K-Nearest Neighbors Classifier with the values of k = 3, 5, and 7 was applied which gave the following results:

S. No.	Value of k	Accuracy Score	Time taken by the algorithm to complete*
1.	3	1.0	0.0094 seconds
2.	5	1.0	0.0091 seconds
3.	7	1.0	0.0092 seconds

*Note: Time taken may vary by 0.0001 – 0.0003 seconds. The behavior, however, is the same. k = 3 takes the longest to complete whereas k = 5 takes the shortest.

3. Decision Tree Classifier

Decision Tree is another non-parametric supervised machine learning technique used for classification and regression. The objective is to create a model that predicts the value of a dependent variable by learning simple decision rules inferred from the data features.

One of the biggest advantages of Decision Trees is their simplicity when it comes to understanding and interpretation. Some of the other advantages are: they can also be visualized; there's no need for much data preparation either---unlike other machine learning techniques---which often require data preprocessing such as data normalization, creation of dummy variables and removal of the blank values. Decision Trees can deal with both categorical and numerical data. However, Decision Tree also has a big disadvantage: over-complex trees might get created that do not generalize the data well. This phenomenon is called overfitting.

The following results were obtained upon the employment of Decision Tree Classifier on to the banknote authentication dataset:

Accuracy Score	0.99
Time taken by the algorithm to complete*	0.0030 seconds

*Note: Time taken may vary by 0.0001 – 0.0004 seconds.

4. Logistic Regression

Logistic regression, despite its name, is a linear model for classification rather than regression. Logistic Regression works well for binary classification. In the case of multi-class classification, logistic regression utilizes a One-vs-Rest or One-vs-All approach for the classification.

After the application of Logistic Regression algorithm on the banknote authentication dataset, following results were obtained:

Accuracy Score	0.99
Time taken by the algorithm to complete*	0.0026 seconds

*Note: Time taken may vary by 0.0001 – 0.0004 seconds.

Performance Comparison

Following is the comparison table between the aforementioned machine learning algorithms when applied to the banknote authentication dataset:

S. No.	Algorithm	Accuracy Score	Time taken by algorithm to complete*
1.	Stochastic Gradient Descent	0.99	0.0027 seconds
2.	K-Nearest Neighbors (for k = 3)	1.00	0.0094 seconds
3.	K-Nearest Neighbors (for k = 5)	1.00	0.0091 seconds
4.	K-Nearest Neighbors (for k = 7)	1.00	0.0092 seconds
5.	Decision Tree Classifier	0.99	0.0030 seconds
6.	Logistic Regression	0.99	0.0026 seconds

*Note: Time taken may vary by 0.0001 – 0.0004 seconds.

Conclusion

Based on the experimental results of the performances of all the above-mentioned machine learning algorithms, K-Nearest Neighbors algorithm with the value of k = 7 seems to be the best-suited algorithm for the banknote authentication dataset because---although it takes longer when compared to the other methods---it also gives the most accurate result.

References

1. Dua, D. and Graff, C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.
2. Chi-Yuan Yeh, Wen-Pin Su, Shie-Jue Lee, Employing multiple-kernel support vector machines for counterfeit banknote recognition, Applied Soft Computing, Volume 11, Issue 1, 2011, Pages 1439-1447, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2010.04.015. (http://www.sciencedirect.com/science/article/pii/S1568494610000918)
3. Syuhada, Nur & Hussin, Burairah & Shibghatullah, A. & Basari, Abd Samad. (2014). BANKNOTE AUTHENTICATION USING ARTIFICIAL NEURAL NETWORK. Science International. 1865-1868.
4. Kaya, Esra & YASAR, Ali & Saritas, Ismail. (2016). Banknote Classification Using Artificial Neural Network Approach. International Journal of Intelligent Systems and Applications in Engineering. 4. 16. 10.18201/ijisae.55250.
5. Shahani, Sumeet & Jagiasi, Aysha & R., Priya. (2018). Analysis of Banknote Authentication System using Machine Learning Techniques. International Journal of Computer Applications. 179. 22-26. 10.5120/ijca2018916343.
6. Ghazvini, Anahita & Awwalu, Jamilu & Abu Bakar, Azuraliza. (2014). Comparative Analysis of Algorithms in Supervised Classification: A Case study of Bank Notes Dataset. International Journal of Computer Trends and Technology. 17. 39-43. 10.14445/22312803/IJCTT-V17P109.
7. Choi, Euisun & Lee, Jongseok & Yoon, Joonhyun. (2006). Feature Extraction for Bank Note Classification Using Wavelet Transform. 2. 934-937. 10.1109/ICPR.2006.553.
8. Scikit-learn: Machine Learning in Python, Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011