

Evaluating Feature Combinations for Banknote Authentication Using Logistic Regression and Gaussian Naive Bayes

Ahmed Ayad Saleh

Department of Computer Science, College of Computer Science and Mathematics,
Tikrit University, Tikrit, Salah al-Din, Iraq
ahmed.ayad@tu.edu.iq

Abstract

The study examines the impact of feature combinations on banknote authentication using two lightweight and interpretable classifiers: Logistic Regression (LR) and Gaussian Naive Bayes (GNB). Experiments were performed on the Banknote Authentication dataset which has 1372 instances and four numerical features: Variance, Skewness, Kurtosis, and Entropy. Two, three, and four features were systematically evaluated in 11 that included combinations. Repeated stratified five-fold cross-validation with 10 repetitions was used to obtain robust estimates of performance. Evaluation measures used were accuracy, precision, recall, F1-Score and ROC-AUC. Logistic Regression performed the best with Variance, Skewness and Kurtosis with accuracy of $98.97\% \pm 0.56\%$, F1-Score of 98.85% and ROC-AUC of 99.98% . This same accuracy and F1-Score was achieved when all four features were used, suggesting that the performance gain of using Entropy was not noticeable for Logistic Regression. Gaussian Naives Bayes was able to get the highest accuracy with $87.22\% \pm 1.60\%$ based on Variance, Skewness and Entropy. The results show that systematic feature-combination analysis is able to find compact representation and at the same time find that the best smaller feature subset varies according to the classifier.

Keywords: Banknote Authentication, Feature Combination, Logistic Regression, Gaussian Naive Bayes, Machine Learning, Classification.

1. Introduction

Automated banknote authentication is an important classification task in which measurable characteristics extracted from banknote images can be used to distinguish between the two target classes [1]. Machine learning provides a practical framework for learning such relationships from labeled examples. The effectiveness of a classifier is influenced by both the learning algorithm and the representation of the input data [2]. Not all the features need to be used, as some can offer little extra information; some combinations can offer a more discriminative representation. A small feature representation can also make a representation more interpretable and decrease unnecessary computational input. The Banknote Authentication dataset is a widely-used benchmark for binary classification. It has 1372 instances and four continuous features derived from the banknote images after transformations (variance, skewness, Kurtosis, entropy) and no missing values, according to the UCI Machine Learning Repository [3].

This research is based on a simple research question, How the combination of features available will impact the performance of two classifiers? Logistic Regression and Gaussian Naive Bayes were selected because they are computationally lightweight and provide complementary modeling assumptions. The unique contribution is to the methodology, which was a systematic comparison of 11 combinations of features, not only with a single train-test split but with repeated stratified cross-validation.

2. Previous Studies

The recent research related to banknote authentication indicates that there are three approaches: feature-based classification on UCI Banknote Authentication dataset, conventional machine-learning classifiers' comparative evaluation, and image/deep-learning algorithms based on country-specific banknote datasets. In the present study, a specific kind of study is considered, including the same UCI dataset, logistic regression, naive bayes and feature/variable selection, as shown in the table below.

Table 1. Summary of previous studies on banknote.

Study	Dataset	Methods	Main relevance
[4] (2023)	UCI Banknote Authentication; 1,372 instances; four features	Logistic Regression, LDA, KNN, SVM, GNB	Systematic variable combinations; LR with Variance + Skewness + Kurtosis reached 0.99 accuracy; entropy was less informative in several combinations.
[5] (2023)	UCI Banknote Authentication	Logistic Regression	Uses the four UCI numerical features and preprocessing/scaling for LR-based authenticity classification.
[6] (2024)	Banknote authentication dataset	KNN, Decision Tree, SVM, Random Forest, Logistic Regression, Naive Bayes, LightGBM	Compares traditional classifiers and LightGBM for counterfeit detection; directly includes LR and Naive Bayes.
[7] (2025)	Banknote Authentication	Multiple machine-learning models	Recent comparative study of machine-learning models for banknote authentication; useful as a current benchmark-oriented reference.
[8] (2025)	UCI benchmark dataset	Logistic Regression, Random Forest, XGBoost	Compares three machine-learning models for banknote forgery detection using a fixed UCI dataset.
[9] (2024)	1,359 images of genuine/counterfeit Iraqi currency	Random Forest, XGBoost, Decision Tree, SVM, CatBoost; VGG16, InceptionV3, MobileNetV2	Provides recent country-specific evidence and a practical image/deep-learning direction for counterfeit-currency detection.
[10] (2024)	Wavelet-based banknote authentication setting	Machine-learning authentication; adversarial analysis	Examines security and robustness of wavelet-based banknote authentication against adversarial manipulation.
[11] (2024)	Image processing and deep learning	Represents the recent shift from handcrafted numerical features toward image-based deep-learning authentication.	Moderate

Study	Dataset	Methods	Main relevance
[12] (2024)	Includes UCI Banknote Authentication (1,372 instances)	Naive Bayes and Bayesian-network classifiers	Uses Banknote Authentication as a benchmark and evaluates probabilistic/Bayesian classification approaches.

3. Evaluating Feature Combinations for Banknote Authentication

3.1 Dataset:

Banknote Authentication dataset was downloaded from UCI Machine Learning Repository [3]. It has 1372 instances, four continuous input features and one integer target variable. The features are Variance, Skewness, Kurtosis and Entropy, table 2 below. The data is extracted from the image that is used in a banknote authentication procedure, the feature values are computed from the image with wavelet transformation, there is no missing value in the repository.

Table 2. Dataset characteristics.

Property	Value
Instances	1,372
Input features	4
Target classes	2
Missing values	None
Feature 1	Variance
Feature 2	Skewness
Feature 3	Kurtosis
Feature 4	Entropy
Class distribution	Class 0: 762; Class 1: 610

3.2 Feature Combination Analysis:

Eleven combinations-two, three, or four features-were tested: six two-feature combinations, four three-feature combinations, and one four-feature combination. Table 3 evaluated feature combinations.

Table 3. Evaluated feature combinations.

Number of Features	Feature Combination
2	Variance + Skewness
2	Variance + Kurtosis
2	Variance + Entropy
2	Skewness + Kurtosis
2	Skewness + Entropy
2	Kurtosis + Entropy
3	Variance + Skewness + Kurtosis
3	Variance + Skewness + Entropy
3	Variance + Kurtosis + Entropy
3	Skewness + Kurtosis + Entropy
4	Variance + Skewness + Kurtosis + Entropy

3.3 Experimental Framework:

The workflow comprised data verification, generation of feature combinations, independent model fitting, stratified five-fold cross-validation, calculation of the metrics, and the selection of the best observed configuration.

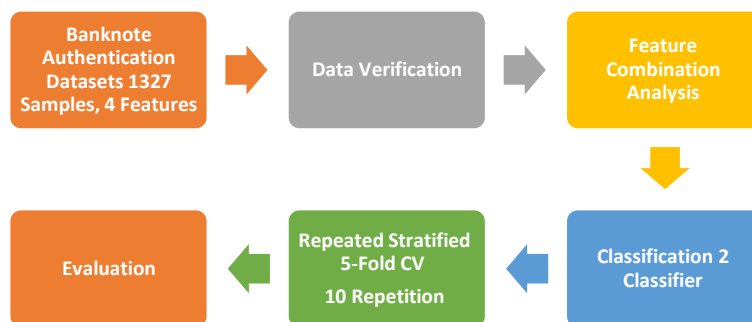


Figure 1. Experimental framework for the feature-combination analysis.

3.4 Classification Algorithms:

This study used two classification algorithms: Logistic Regression (LR) and Gaussian Naive Bayes (GNB) to assess the classification performance. These algorithms were chosen because they are simple and efficient to run and appropriate for the evaluation of the selected feature combinations. Here are brief descriptions of the two classifiers.

3.4.1 Logistic Regression:

Logistic Regression is a supervised binary classification model which estimates class probability by logistic function. The simplicity of its structure allows it to be analyzed numerically, with an interpretation that is understandable.

$$P(y=1|X) = 1 / (1 + e^{(-z)}), \text{ where } z = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n. [13]$$

3.4.2 Gaussian Naive Bayes:

The Gaussian Naive Bayes is a probabilistic classifier using the Bayes' theorem and the assumption that features are all conditionally independent, given the class. Gaussian distributions are used to model continuous variables.

$$P(C_k|X) \propto P(C_k) \prod_i P(x_i|C_k) [14].$$

3.5 Repeated Stratified Five-Fold Cross-Validation:

Stratified five-fold cross validation (with 10 repetitions) was used. The data were split into 5 folds in each repetition with a similar class distribution. Four folds were used for training and one fold for

validation, and this was done until all four folds had been used as the validation fold. All the repeated results were averaged and the standard deviation computed.

3.6 Evaluation Metrics:

Accuracy, Precision, Recall, F1-Score and ROC-AUC were used. The overall correctness is measured by Accuracy, the number of correctly predicted positive cases is measured by Precision, the number of positive cases detected is measured by Recall, the harmonic mean of Precision and Recall is measured by F1-Score, and the sum of the discrimination across thresholds is measured by ROC-AUC.

4. Results

A summary of the experimental results of the classifiers tested is presented in Table 4. Best observed performance of each classifier with their optimal configuration is presented.

Table 4. Best observed performance for each classifier.

Model	Best Features	Accuracy	Precision	Recall	F1	ROC-AUC
Logistic Regression	Variance + Skewness + Kurtosis	0.9897 ± 0.0056	0.9839	0.9933	0.9885	0.9998
Gaussian Naive Bayes	Variance + Skewness + Entropy	0.8722 ± 0.0160	0.8696	0.8390	0.8535	0.9526

4.1 Accuracy Versus Number of Features:

The best accuracies achieved at each feature-count level are presented in Figure 2. As the number of features is increased, Logistic Regression improved from 0.8824 for 2 features to 0.9897 for 3 features, but the fourth feature did not facilitate any further progress. Gaussian Naive Bayes achieved approximately 0.8722 with two and three features, but decreased to 0.8396 with four features.

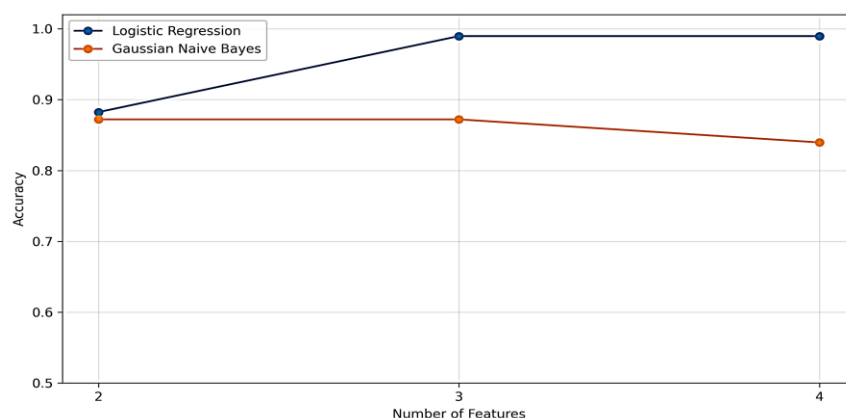


Figure 2. Accuracy versus number of features.

4.2 F1-Score Versus Number of Features:

As per figure 3, Logistic Regression achieved an F1-Score of 0.9885 using 3 features and it was achieved using 4 features. The Gaussian Naives Bayes model scored the highest F1-Score (0.8535) with only three features.

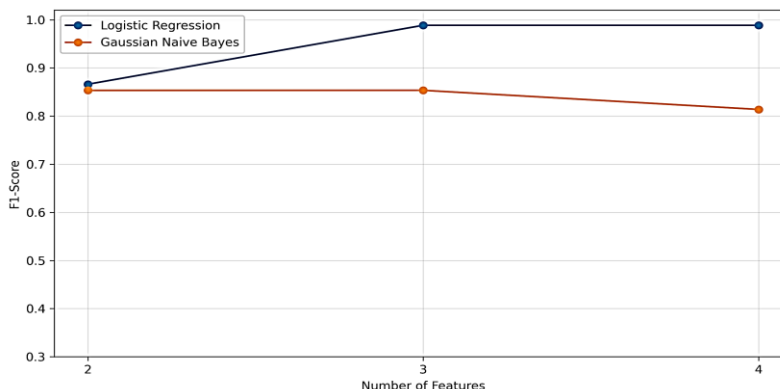


Figure 3. F1-Score versus number of features.

4.3 Logistic Regression Feature Combination Analysis:

Logistic Regression, the best compact representation was found to be Variance + Skewness + Kurtosis, with the best performance: Accuracy = 0.9897, F1-Score = 0.9885, and ROC-AUC = 0.9998. The same result was obtained for the four-feature combination, with respect to both the Accuracy and F1-Score. Thus, Entropy did not provide an observable improvement for this classifier.

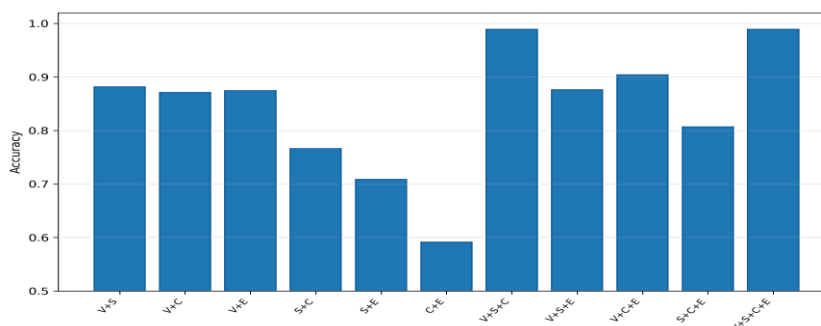


Figure 4. Logistic Regression performance across feature combinations.

4.4 Gaussian Naive Bayes Feature Combination Analysis:

Variance + Skewness + Entropy had the highest accuracy (0.8722) and F1-Score (0.8535) for Gaussian Naives Bayes. The two feature combination Variance + Skewness came close to the same accuracy. Figure 5.

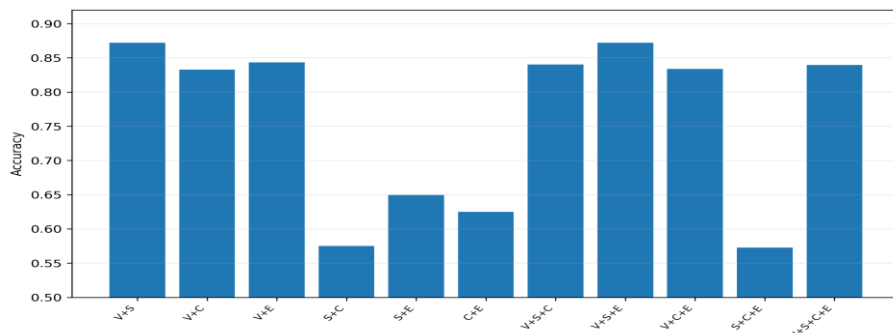


Figure 5. Gaussian Naive Bayes performance across feature combinations.

4.5 Final Performance Comparison:

The best configuration of each of the classifiers is compared in Figure 6 with respect to the five evaluation metrics. Logistic Regression had the best results on all reported measures. The greatest accuracy difference between the two optimum configurations was about 11.75 percentage points.

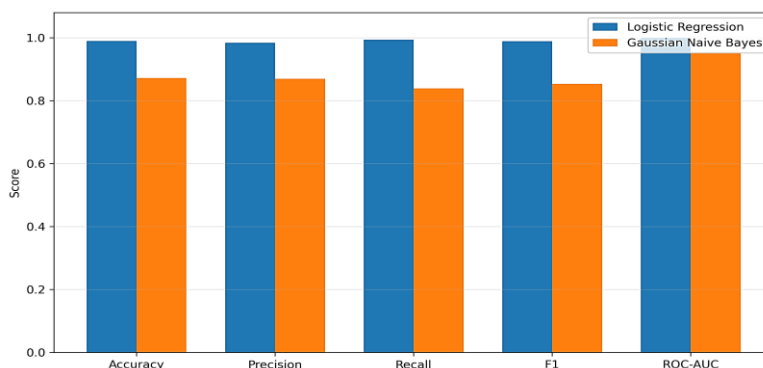


Figure 6. Final performance comparison using the best feature combination for each classifier.

5. Comparison with Related Study

The results from our current study should be compared to the results of a previous study. The study to be compared is considered to be the closest methodological match, Rick Nie (2023), it is based on the same UCI Banknote Authentication dataset, on the same four numerical features and explicitly addresses the classifiers Logistic Regression and Gaussian Naive Bayes. As can be seen in the table below, the combination of three features (Variance + Skewness + Kurtosis) gave the best result for Logistic Regression with 0.99 accuracy, 0.98 precision, 0.99 recall, and 0.99 F1--score for Nie's experiment.

Table 5. Performance comparison with previous study.

Study	Best feature combination	Accuracy	Precision	Recall	F1-score	Evaluation protocol
Our study Logistic Regression	Variance + Skewness + Kurtosis	98.97% ± 0.56%	98.39%	99.33%	98.85%	Repeated stratified 5-fold CV × 10
Nie (2023) Logistic Regression	Variance + Skewness + Kurtosis	99.00%	98.00%	99.00%	99.00%	80:20 train/test split
Our study Gaussian Naive Bayes	Variance + Skewness + Entropy	87.22% ± 1.60%	86.96%	83.90%	85.35%	Repeated stratified 5-fold CV × 10

The Logistic Regression result of the present study (98.97 % accuracy) is very close to the accuracy obtained by Nie (2023) (99.00 % accuracy) for the same three feature combination. This agreement further confirms the finding that Variance, Skewness and Kurtosis are a very informative compact representation in Logistic Regression on the UCI Banknote Authentication dataset. The present study, however, employs repeated cross-validation and hence a standard deviation is also reported. The present study finds that, for Gaussian Naive Bayes, the optimal feature subset is different, namely Variance + Skewness + Entropy, which shows that the optimal features combination is different for different classifiers.

6. Discussion

The results reveal significant impact of feature combination on classification performance and their impact varies with the classifiers. The best result was Logistic Regression with 98.97% accuracy with just three features: Variance, Skewness, and Kurtosis. And the complete four-feature representation did not affect the accuracy and F1-Score. Thus, the omission of Entropy is not affecting the main performance measures for the tested configurations, and does not appear to be necessary. Variance + Skewness + Entropy was the feature set that gave the best accuracy for Gaussian Naive Bayes, whereas adding the Kurtosis gave a slight decrease in accuracy. This illustrates that increase in variables does not necessarily lead to better classification.

Lower accuracy variability, or the standard deviation across repeated-CV (0.0056 for the best Logistic Regression configuration, and 0.0160 for the best Gaussian Naive Bayes configuration), was also found for the Logistic Regression configuration in this experiment. Conclusions cannot be transferred to other banknote datasets without external validation and are limited to the banknote dataset, models and experimental protocol used. It is also in agreement with the work by Nie (2023) which used the same Banknote Authentication dataset and found that the combination of Variance, Skewness, and Kurtosis is the best set of features for Logistic Regression. The accuracy obtained in the present study (98.97%) was very close to the 99.00% reported by Nie (2023). While the evaluation protocols vary, this agreement adds support to the effectiveness of this representation with Logistic Regression.

7. Conclusion

This paper compared 11 combinations of features to authenticate banknotes with Logistic Regression and Gaussian Naive Bayes using repeated stratified five-fold cross-validation with 10 runs. Logistic Regression achieved the best overall performance using Variance, Skewness, and Kurtosis, with Accuracy = $98.97\% \pm 0.56\%$, Precision = 98.39%, Recall = 99.33%, F1-Score = 98.85%, and ROC-AUC = 99.98%. The same Accuracy and F1-Score was achieved with all four features, which means that Entropy was not required to keep the same performance as observed. Variance, Skewness, and Entropy were the best values for Gaussian Na Bayes with an accuracy of $87.22\% \pm 1.60\%$. In conclusion, the study illustrates that a systematic feature-combination analysis is a very simple and interpretable method for finding compact representations in low dimensional classification problems.

8. Limitations and Future Work

The analysis of this study is performed on a single publicly available benchmark dataset with four pre-extracted numerical features. Future work can explore automated feature-selection methods, additional classifiers, model calibration and external validation on independent banknote datasets.

References

- [1] F. Van Der Horst, J. Snell, and J. Theeuwes, “Enhancing banknote authentication by guiding attention to security features and manipulating prevalence expectancy,” *Cogn. Res. Princ. Implic.*, 2021, doi: 10.1186/s41235-021-00341-x.
- [2] N. G. Abadi, “Machine Learning-Based Authentication of Banknotes : A Comprehensive Analysis,” vol. 4, no. 1, pp. 22–30, 2024.
- [3] V. Lohweg, “Banknote Authentication,” UCI Machine Learning Repository. [Online]. Available: <https://archive.ics.uci.edu/dataset/267/banknote+authentication>
- [4] R. Nie, “Comparative Analysis Of Machine Learning Based Bank Note Authentication Through Variable Selection,” pp. 1–12, 2023.
- [5] P. A. Singh, P. Z. Hasan, and P. S. Sharma, “Improving Banknote Authentication Accuracy through Logistic Regression and Data Preprocessing,” vol. 10, no. 1, 2022, doi: 10.15680/IJIRCCE.2022.1001044.
- [6] K. Sn and K. Chitra, “Evaluating Machine Learning Algorithms for Fake Currency Detection,” vol. 2024, no. June, 2024.
- [7] J. J. Jena, “A Comprehensive Exploration of Machine Learning Models for Banknote Authentication,” *2025 Int. Conf. Intell. Cloud Comput.*, pp. 1–6, 2025, doi: 10.1109/ICoICC64033.2025.11052182.

-
- [8] O. V. Eşİdır, G. Bak, and B. Karadağ, “Finansal Ve Medyatik Sahteciliğe Karşı Yapay Zekâ Çözümleri : Banknot Sahteciliği Tespitinde Makine Öğrenmesi Uygulamaları Artificial Intelligence Solutions Against Financial And Media Counterfeiting : Machine Learning Applications İn Banknote Counterfeit Detection,” vol. 17, no. 33, pp. 1066–1076, 2025.
- [9] N. M. Z. Hamed and F. Al Azzo, “Develop a Robust System for Detecting Counterfeit Iraq Currencies Based on Deep Learning Techniques,” vol. 33, no. 2, 2024.
- [10] J. K. C.-A. H. V. Lohweg, “Hidden in Plain Sight: Adversarial Attack on Wavelet-Based Banknote Authentication,” *Res. Technol. Soc. Ind. Innov.*, 2024, doi: 10.1109/RTSI61910.2024.10761182.
- [11] T. F. Document, “Identification of Counterfeit Currency Using Image Processing and Deep Learning,” *IEEE 5th Int. Conf. Conver. Technol.*, 2019, doi: 10.1109/I2CT45611.2019.9033740.
- [12] S. Sugahara and J. Cussens, “Learning Bayesian Network Classifiers to Minimize Class Variable Parameters,” vol. 27, pp. 1–41, 2026.
- [13] A. Bhandare, S. Biradar, N. Lonari, and V. Pawar, “Online Subsidy Management System Using Machine Learning (Algorithm- Logistic Regression , Random Forest , Decision Tree),” vol. 3, no. 4, pp. 17–21, 2026, doi: 10.5281/zenodo.19899201.
- [14] M. V. Anand, B. Kiranbala, S. R. Srividhya, C. Kavitha, and M. Younus, “Gaussian Naïve Bayes Algorithm : A Reliable Technique Involved in the Assortment of the Segregation in Cancer,” vol. 2022, 2022, doi: 10.1155/2022/2436946.