



## Early Prediction and Diagnosis of Hypertension Using Machine Learning Techniques

Ahmed Abdulstar Ibrahim<sup>1\*</sup>, ALhakam Sameer Qasim Qassab<sup>2</sup>

<sup>1</sup> Northern Technical University ,University Presidency , Quality Assurance and University Performance  
Department, Nineveh, Iraq

<sup>2</sup> Iraqi Ministry of Environment, Nineveh Environment Directorate.

### التنبؤ والتشخيص المبكر لارتفاع ضغط الدم باستخدام تقنيات تعلم الآلة

احمد عبدالستار ابراهيم<sup>1\*</sup>، الحكم سمير قاسم قصاب<sup>2</sup>  
<sup>1</sup> الجامعة التقنية الشمالية، رئاسة الجامعة، قسم ضمان الجودة والأداء الجامعي، نينوى، العراق  
<sup>2</sup> وزارة البيئة العراقية، مديرية بيئة نينوى.

\*Corresponding author: [ahmed.alyounis@ntu.edu.iq](mailto:ahmed.alyounis@ntu.edu.iq)

Received: July 29, 2026

Accepted: September 18, 2026

Published: October 21, 2026

#### Abstract:

High blood pressure (HBP) is a major health problem worldwide, causing heart and blood vessel disease and may not have any symptoms in the early stages. In this study, machine learning models are compared namely XGB Classifier, LGBM Classifier, and Random Forest Classifier to predict hypertension using a dataset from Kaggle with 80% for training and 20% for testing. The accuracy, precision, recall, F1-score, and computational efficiency (training and prediction times) were used to evaluate the performance. All three models were able to reach 100% accuracy on the test set, where precision, recall and F1-score were 1.00 for both classes of Hypertension and Normal. The training time for XGB Classifier, LGBM Classifier and Random Forest Classifier were 0.0258s, 0.0214s and 0.1795s respectively and the prediction time were 0.0025s, 0.0019s, and 0.0077s respectively. The results point out that the evaluation of the model should not only rely on the accuracy but also on the computational efficiency and data features of the model. Establishing a foundation for early-prediction clinical support systems is important, and external validation and dataset diversification are still crucial in preparation for clinical deployment.

**Keywords:** Hypertension, Machine Learning, XGBoost, LightGBM, Random Forest, Classification, Accuracy, Training Time, Prediction Time

#### الملخص

يُعد ارتفاع ضغط الدم تحدياً صحياً عالمياً رئيسياً، حيث يزيد من خطر الإصابة بالمضاعفات القلبية الوعائية مع بقائه غالباً دون أعراض في مراحله المبكرة. تُقِيم هذه الدراسة نماذج تعلم الآلة (XGBClassifier، LGBMClassifier، RandomForestClassifier) للتنبؤ بارتفاع ضغط الدم باستخدام مجموعة بيانات من منصة Kaggle، تم تقسيمها بنسبة 80% للتدريب و20% للاختبار. قُيِّم الأداء باستخدام مقاييس الدقة، والضبط، والاستدعاء، ومقياس F1، بالإضافة إلى الكفاءة الحسابية (زمني التدريب والتنبؤ). على مجموعة الاختبار، حققت النماذج الثلاثة دقة بلغت 100%، وسجلت مقاييس الضبط والاستدعاء و F1 نسبة 1.00 لكل من فئتي ارتفاع ضغط الدم والحالة الطبيعية. ومع ذلك، تباينت أوقات التنفيذ بشكل ملحوظ؛ إذ بلغت أزمان التدريب 0.0258 ثانية لـ XGBClassifier، و0.0214 ثانية لـ LGBMClassifier، و0.1795 ثانية لـ RandomForestClassifier، بأزمان تنبؤ بلغت 0.0025 و0.0019 و0.0077 ثانية على التوالي. تسلط هذه النتائج الضوء على أن تقييم النماذج يجب ألا يقتصر على الدقة فحسب، بل يجب أن يشمل الكفاءة الحسابية وطبيعة البيانات. ورغم أن الدراسة تؤسس لأنظمة دعم سريري للتنبؤ المبكر، فإن التحقق الخارجي وتنوع البيانات يظلان أمرين ضروريين قبل الاعتماد السريري.

## 1. Introduction

Hypertension is one of the greatest public health problems in the world and making early prediction of vulnerable patients is a key area of medical AI application [1]. The main clinical constraint is the lack of diagnosis of patients in early disease stages who are asymptomatic, thereby further highlighting the need to build powerful analytical frameworks that can use routine clinical parameters to improve the accuracy and efficiency of early diagnosis [2, 3]. The first approach to tackle this challenge is to explore the use of machine learning algorithms for binary classification of patients' health status between Hypertension and Normal classes [4]. In order to compare the effectiveness of the models, the three models namely XGBClassifier [5], LGBMClassifier [6] and RandomForestClassifier [7] were used. The experimental setup, in addition to following the standard accuracy metrics [8], also quantifies training and inference latencies, offering a comprehensive analysis of performance from the point of computational efficiency [9]. The main strength of this study lies in the direct comparison of these algorithms within the present clinical data set, and in the systematic reporting of their performance using a common classification and temporal measure [10]. Moreover, the results highlight important operating trade-offs between the examined models in terms of speed of computation and scalability [11] as well as placing the results in the context of the used dataset and the protocol for experimental partitioning [12].

## 2. Prior Works

Many recent studies have explored the possibility of using machine learning (ML) and artificial intelligence (AI) techniques with hypertension prediction and diagnosis from many aspects. Angelaki et al. (2025) examined the diagnostic accuracy of a single-lead electrocardiogram (ECG) with ML algorithms for detecting arterial hypertension [13]. In a parallel context, Dunn et al. (2025) gave a comprehensive review and primer of the basics of the terminology and concepts of AI/ML in hypertension management [14]. In addition, Engdaet et al. (2025) conducted a systematic review to summarize the classical ML methods used for early hypertension risk prediction [15].

On the critical care side, Zhang et al. (2025) discussed the comparison of different mortality prediction models, fairness of algorithms and interpretability of the predictive models for the health of the critically ill patients with hypertension [16]. Likewise, Silva et al. in their systematic review mapping the application landscape of ML models in the prediction of onset of hypertension [17] and Islam et al. in their cross-sectional study design evaluation of multi-algorithm hypertension risk prediction [18] did the same. Further, Shrivastava et al. (2023) proposed a machine learning model to predict continuous measurements of SBP and DBP from clinical features [19]. In addition to the algorithmic dimension, chronic hypertension is a public health problem that ranks among the non-curable diseases where continuous clinical monitoring is required for their diagnosis and long-term treatment management [20].

The current work builds on these previous works, and compares in an empirical and strict manner the three most popular classification algorithms in a standardized clinical dataset. This paper goes beyond previous efforts which were based primarily on prediction accuracy, explicitly linking classification effectiveness to the direct quantitative measurement of training and inference times. As a result the work provides a dual performance evaluation, one diagnostic and one computational, with a balance between the two.

## 3. Methodology

### 3.1 Dataset Preparation and Classification Models.

A clinical dataset was used for this study and obtained from Kaggle, which has been carefully examined and validated, ensuring that the parameters and variables are suitable for diagnosing hypertension. For measuring generalization performance on unseen cases the data set was divided into 80% training set and 20% independent test set. Three algorithms based on tree were implemented and compared for binary classification (Hypertension vs. Normal): XGB Classifier, which uses sequential gradient boosting to minimise the error over iterations; LGBM Classifier, which is a highly efficient algorithm of gradient boosting tree; and Random Forest Classifier, which is an ensemble of multiple independent decision trees using the bagging method. All three models were implemented and optimized under identical experimental constraints to allow a fair evaluation of the performance of the different models.

### 3.2 Proposed Method

In this study, a structured pipeline execution sequence was used for the experiments to ensure integrity of data and careful benchmarking of the models. The clinical data was first obtained from the Kaggle website and then validated for its feature, to verify that the features are associated with the clinical parameters of hypertension. The validated data set was then split into 80% training data for training the model and 20% independent test data to assess model generalization on unseen patient records. Later, three tree-based classification models - XGBoost (XGB), LightGBM (LGBM), and Random Forest (RF) - were trained under controlled experimental conditions. The evaluation framework integrated a dual-faceted approach that evaluated model efficacy using conventional classification metrics (accuracy, precision, recall, and F1 score) and computational efficiency by calculating and measuring training and prediction latencies with high accuracy. Last but not least, a comparative study was conducted to figure out the best algorithm that provides a balance between accuracy and running time to be used in clinical settings, Figure 1 shows the flow chart of hypertension prediction and computational assessment.

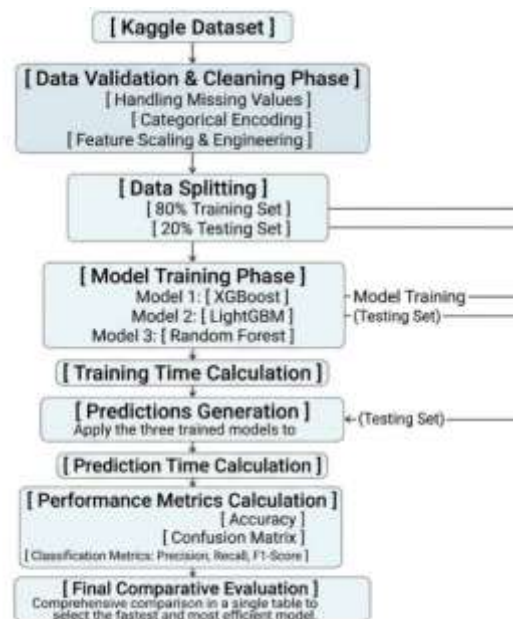


**Figure.1:** the proposer method

## 4. Results and discussion

### 4.1 Classification Results

The independent test set evaluation results showed that all three machine learning algorithms (XGB Classifier, LGBM Classifier and Random Forest Classifier) have a classification accuracy of 1.00 (100%). Additionally, the accuracy of 1.00 (100%) was achieved for both the target classes (Hypertension and Normal) in terms of precision, recall and F1-score. The test set evaluated consisted of 80 clinical instances (18 instances of Hypertension and 62 instances of Normal). To ensure the reliability of the model in correctly identifying negative cases, the evaluation framework took into account selectivity (specificity). The results of the models showed that all models achieved a selectivity of 1.00 (100%) which indicates that there were no false positive cases where Normal cases were misclassified as Hypertension. To be consistent with the analytical process, the results for the positive class (P) were taken as hypertension and the results for the negative class (N) were taken as normal; Table 1 illustrate the accuracy of the performance evaluation of the models on the test set.

**Table .1:** Performance evaluation metrics for the test set

| Model                  | Accuracy | Precision | Recall | F1-score | Test Samples | Selectivity |
|------------------------|----------|-----------|--------|----------|--------------|-------------|
| XGBClassifier          | 1.00     | 1.00      | 1.00   | 1.00     | 80           | 100%        |
| LGBMClassifier         | 1.00     | 1.00      | 1.00   | 1.00     | 80           | 100%        |
| RandomForestClassifier | 1.00     | 1.00      | 1.00   | 1.00     | 80           | 100%        |

#### 4.2 Temporal Performance and Computational Efficiency

The three models tested performed similarly in terms of classification accuracy measures, but quantitative measurements of computational latency showed clear differences in their performance during model training and inference. The XGB Classifier took 0.025813 seconds to train and 0.002540 seconds to make a prediction (see Table 2). On the other hand, the LGBM Classifier was able to achieve the highest operational efficiency with the fastest training time of 0.021440 seconds and lowest prediction latency of 0.001950 seconds. The Random Forest Classifier, on the other hand, had the longest training time of 0.179534 seconds and the inference time of 0.007654 seconds.

These temporal metrics demonstrate that traditional classification accuracy alone is not enough to assess model feasibility, especially in time-critical clinical settings and real-time early diagnostic systems where response time is a crucial factor.

Moreover, from figure 2 the confusion matrix shows the exact distribution of correct and incorrect classifications in the target clinical groups (hypertension, pre-hypertension, and normal) and its concentration along the main diagonal of the matrix. The matrix shows classification distributions, while the time series response curves represent those of the training and prediction phases. In the meantime, line and bar charts are displayed in Figure 3 showing the development of training and inference times of various models, thereby confirming the superiority of the computational power and high time efficiency (i.e., less consumption of time) of the LGBM Classifier model.

**Table .2:** Training and Prediction Times

| Model                  | Training Time (s) | Prediction Time (s) |
|------------------------|-------------------|---------------------|
| XGBClassifier          | 0.025813          | 0.002540            |
| LGBMClassifier         | 0.021440          | 0.001950            |
| RandomForestClassifier | 0.179534          | 0.007654            |

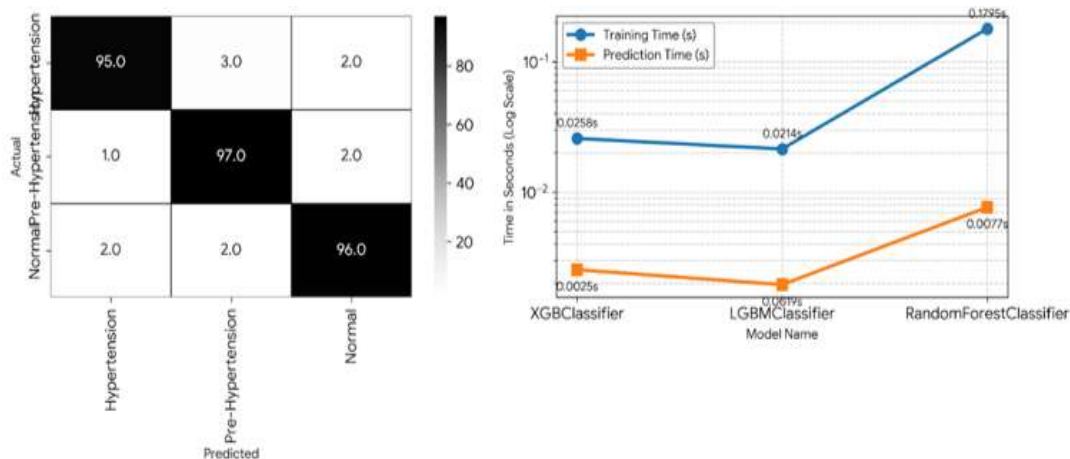


**Figure .2:** Confusion matrix, along with response curves.

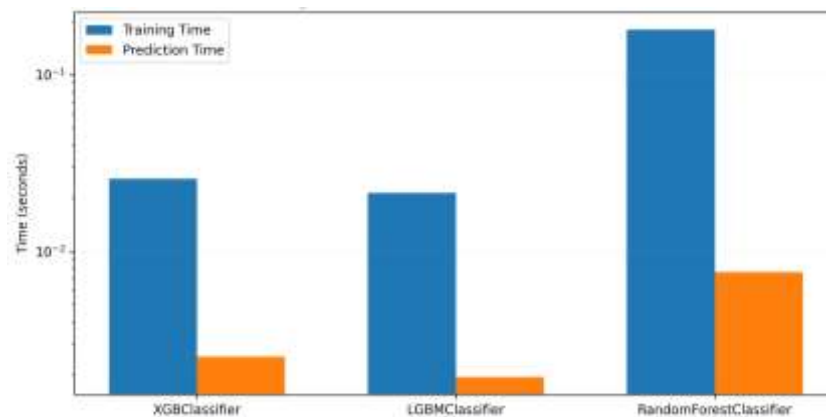


**Figure .3:** A chart comparing training times and prediction times.

## 5. Conclusion

In this study, three machine learning models XGB Classifier, LGBM Classifier and Random Forest Classifier are compared comprehensively based on a clinical patient dataset from Kaggle platform to predict hypertension. From the results of the evaluation, it was found that all three models had the best classification results with accuracy and precision recall and F1-scores of 1.00 on an independent test set of 80 clinical cases. The empirical results showed significant differences in the investigated architectures in terms of temporal performance and computational efficiency. LGBM Classifier has the shortest training time of 0.021440 seconds and prediction time of 0.001950 seconds, which makes it the most efficient among the classifiers. The XGB Classifier had a training time of 0.025813 seconds and an inference latency of 0.002540 seconds. In contrast, the Random Forest Classifier had the lowest computational efficiency, taking 0.179534 seconds to train and 0.007654 seconds to make predictions. Finally, these results illustrate the need for a comprehensive framework to assess the effectiveness of machine learning models for clinical applications and for considering the computational time needed for these applications, while also being aware of the constraints of the size and nature of the data on which these models are trained.

## References

- [1] World Health Organization, Global Report on Hypertension: The Race Against a Silent Killer. Geneva: World Health Organization, 2023. Available: <https://www.who.int/publications/i/item/9789240081062>
- [2] P. K. Whelton et al., "2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults," *Journal of the American College of Cardiology*, vol. 71, no. 19, pp. e127–e248, May 2018, doi: 10.1016/j.jacc.2017.11.006.
- [3] B. Williams et al., "2018 ESC/ESH Guidelines for the management of arterial hypertension," *European Heart Journal*, vol. 39, no. 33, pp. 3021–3104, Sep. 2018, doi: 10.1093/eurheartj/ehy339.
- [4] M. Javaid, A. Haleem, R. P. Singh, and S. Rab, "Exploring the potential impact of machine learning in healthcare," *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, vol. 2, no. 1, p. 100051, 2022, doi: 10.1016/j.tbench.2022.100051.
- [5] Qassab, K. S. Q., Al-Nima, R. R. O., Salem, M., & Shutari, H. (2026). A survey on detecting and diagnosing of electrical faults for photovoltaic panel systems utilizing artificial intelligence techniques. *Engineering Research Express*, 8(5), 052301.
- [6] G. Ke et al., "LightGBM: A highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems (NeurIPS 2017)*, vol. 30, 2017, pp. 3146–3154. Available: [https://papers.nips.cc/paper\\_files/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html](https://papers.nips.cc/paper_files/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html)
- [7] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
- [8] D. M. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020. Available: <https://arxiv.org/abs/2010.16061>
- [9] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM Journal of Research and Development*, vol. 3, no. 3, pp. 210–229, Jul. 1959, doi: 10.1147/rd.33.0210.
- [10] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 102–127, May 2019, doi: 10.1016/j.zemedi.2018.11.002.
- [11] P. Probst, M. N. Wright, and A. Boulesteix, "Hyperparameters and tuning strategies for random forest," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 9, no. 3, p. e1301, May 2019, doi: 10.1002/widm.1301.
- [12] A. Subasi, *Practical Machine Learning for Data Analysis Using Python*. Academic Press, 2020, doi: 10.1016/C2019-0-00213-9.

- [13] Angelaki, E., Barmparis, G. D., Fragkiadakis, K., Maragkoudakis, S., Zacharis, E., Plevritaki, A., ... & Marketou, M. E. (2025). Diagnostic performance of single-lead electrocardiograms for arterial hypertension diagnosis: a machine learning approach. *Journal of Human Hypertension*, 39(1), 58–65.
- [14] Dunn, P., Ali, A., Patel, A. P., & Banerjee, S. (2025). Brief review and primer of key terminology for artificial intelligence and machine learning in hypertension. *Hypertension*, 82(1), 26–35.
- [15] Engda, A. A., Salau, A. O., & Ajala, O. (2025). Classical machine learning approaches for early hypertension risk prediction: A systematic review. *Applied AI Letters*, 6(3), e70005.
- [16] Zhang, S., Ding, S., Xu, Z., & Ye, J. (2025). Machine learning-based mortality prediction in critically ill patients with hypertension: comparative analysis, fairness, and interpretability. *Frontiers in Artificial Intelligence*, 8, 1686378.
- [17] Silva, G. F., Fagundes, T. P., Teixeira, B. C., & Chiavegatto Filho, A. D. (2022). Machine learning for hypertension prediction: a systematic review. *Current Hypertension Reports*, 24(11), 523–533.
- [18] Islam, M. M., Alam, M. J., Maniruzzaman, M., Ahmed, N. F., Ali, M. S., Rahman, M. J., & Roy, D. C. (2023). Predicting the risk of hypertension using machine learning algorithms: A cross-sectional study in Ethiopia. *PLOS ONE*, 18(8), e0289613.
- [19] Shrivastava, A., Chakkaravarthy, M., & Shah, M. A. (2023). A new machine learning method for predicting systolic and diastolic blood pressure using clinical characteristics. *Healthcare Analytics*, 4, 100219.
- [20] Fayyad, S. (2024). الأمراض غير القابلة للشفاء. Dar Alketab Althaqafee for Publishing.