Al-Ghary Journal of Economic and Administrative Sciences Vol. 22 (No.2) 2026 PP. 131- 150

مجلة الغري للعلوم الاقتصادية والادارية
مجلد (22) عدد (2) 2026

The Impact of Multiple Comorbidities on ST-Segment Abnormalities in Electrocardiogram (ECG) Signals with Machine Learning

Sara Noori Mohammad Ali
University of Sulaimani, College of Administration and Economics
sara.mohammad@univsul.edu.iq

Nawzad Muhammed Ahmed
University of Sulaimani, College of Administration and Economics
nawzad.mahmud@univsul.ed.iq

Abstract

In this work, a rigorous comparative analysis of five machine learning to choose which of them best ensemble models are AdaBoost, GB, CatBoost, XGBoost, and LightGB, was performed on four varying data partitioning schemes (from 70-30 to 85-15 train-test splits) for maximizing clinical risk stratification. The sample consisted of (1000) ST-segment abnormal patients, received from Shar Hospital in Sulaymaniyah City's. The dataset included age, gender, smoking status, hypertension, chronic cardiovascular disease, diabetes mellitus, dyslipidemia, and hypothyroidism. Chi-square test revealed substantial correlations among the predictor variables. The five boosting models were subsequently trained and tested under a range of data-splitting regimes, with performance metrics in terms of accuracy, F1-score, recall, and precision. Computational protocols were executed by using Python programming **Results**: by using Chi-square test get significant correlation between ST-segment and each of Age, Hypertension, Chronic Cardiovascular and Dyslipidemia. After that best Algorithm is AdaBoost performed best discriminatively, with a highly good F1-score of (0.8839) and maximum flawless recall (1.000) at (75-25) train-test split in all experiments. Feature Importance: Age was the strongest predictor, which reaffirms its prognostic value in ST-segment related pathology.

Keywords: ***(****ST-segment,Chi-square test,AdaBoost,GB, CatBoost, XGBoost, LightGB****)***

Al-Ghary Journal of Economic and Administrative Sciences Vol. 22 (No.2) 2026 PP. 132- 150

مجلة الغري للعلوم الاقتصادية والادارية
مجلد (22) عدد (2) 2026

تأثير الأمراض المصاحبة المتعددة على تشوهات القطعة ST في إشارات تخطيط كهربية القلب (ECG) باستخدام التعلم الآلي

سارا نورى محمد على
جامعه السليمانيه ، كليه الادارة والاقتصاد
sara.mohammad@univsul.edu.iq

نوزاد محمد احمد
جامعه السليمانيه ، كليه الادارة والاقتصاد
nawzad.mahmud@univsul.ed.iq

المستخلص

في هذا العمل، تم إجراء تحليل مقارن دقيق لخمسة نماذج تعلم آلي لاختيار أيهما أفضل من نوع "التجميع (Ensemble Models) وهي (XGBoost, CatBoost, GB, AdaBoost , LightGB) باستخدام أربعة أنماط لتقسيم البيانات (من 70-30 إلى 85-15 لتقسيم التدريب-الاختبار) لتعظيم التقييم السريري للمخاطر. اخذ العينة (1000) مريض يعانون من تشوهات في قطعةST مأخوذة من مستشفى شار في مدينة السليمانية. تضمنت البيانات (العمر، الجنس، التدخين، ارتفاع ضغط الدم، أمراض القلب المزمنة ، السكري، اختلال الدهون وقصور الغدة الدرقية). كشف التحليل الإحصائي الأولي باستخدام اختبار مربع كاي (Chi-square) عن وجود علاقات قوية بين بعض المتغيرات. بعد ذلك، تم تدريب النماذج الخمسة واختبارها تحت أنظمة تقسيم مختلفة، مع قياس الأداء باستخدام الدقة(accuracy) ، ودرجةF1 ، والاستدعاء(recall) ، والدقة التنبؤية(Precision) . باستخدام الحاسوب (Python programming) **النتائج:** أظهر اختبار مربع كاي ارتباطًا كبيرًا بين تغيرات قطعة ST وكل من :العمر، ارتفاع ضغط الدم، أمراض القلب المزمنة، واختلال الدهون. كان أفضل خوارزمية أداءً AdaBoost حيث حقق أعلى درجة F1 (0.8839) واستدعاء مثالي (1.000) عند تقسيم (75-25) ، العمر كان العامل الأكثر تأثيرا.

الكلمات المفتاحية *(القطعة ST، اختبار مربع كاي، AdaBoost، GB، CatBoost، XGBoost، LightGBM)*

1- Introduction

Cardiovascular diseases (CVDs) persist as the predominant cause of global mortality, underscoring persistent gaps in preventive strategies, early diagnosis, and equitable healthcare access (Jurado et al., 2022; Mehrang et al., 2018). Myocardial infarction (MI), a critical CVD manifestation, exhibits rapid progression and elevated fatality rates, necessitating urgent intervention (Chen et al., 2025). In emergency settings, the electrocardiogram (ECG) serves as the gold-standard diagnostic tool, offering real-time, non-invasive assessment of cardiac electrophysiology to guide time-sensitive management of acute coronary syndromes and arrhythmias (Liu et al., 2025; Al Younis et al., 2024). The ECG enables detection of ischemic and structural cardiac pathologies, proving indispensable for clinical decision-making (Mamun & Alouani, 2020; Kalmady et al., 2024).

ST-segment elevation myocardial infarction (STEMI), the most critical form of acute coronary syndrome, requires immediate ECG diagnosis and emergency reperfusion therapy to prevent irreversible heart muscle damage and reduce death rates. Time-to-treatment is crucial for patient survival and recovery. (Vogel et al., 2019). According to the World Health Organization, an estimated 7.3 million deaths annually are attributed to acute heart attacks, underscoring the need for rapid and accurate ECG interpretation (Bodini et al., 2020). ST-segment deviations (whether elevation or depression) play a pivotal role in distinguishing between STEMI and non-ST-segment elevation MI (NSTEMI), with STEMI indicating full thickness myocardial injury and NSTEMI involving partial-thickness ischemia (Symposium, 2024; Al Younis et al., 2024). Given the critical role of ECG in early MI detection, advancements in automated and real-time analysis could significantly enhance diagnostic precision and reduce time to treatment, ultimately improving survival rates and long term cardiac health.

2- Working methods

ST-segment elevations (STE) or depressions (STD) are characteristic ECG presentations of acute coronary syndromes. The distribution of demographic and clinical characteristics among patients presenting with STE versus STD was analyzed using Chi-square tests of independence. The Figure (1) show the ST-segment on ECG. Machine learning (ML) in the medical field has been predominantly focused on classification tasks, particularly in the analysis of medical images. (Nayak et al., 2021). ML employs diverse statistical techniques and algorithms to enable computational systems to extract knowledge from data and improve decision making processes autonomously, without relying on explicit programming. In healthcare applications, this capability allows for enhanced diagnostic accuracy, treatment optimization, and operational efficiency through data-driven insights. (Kasim et al., 2024)

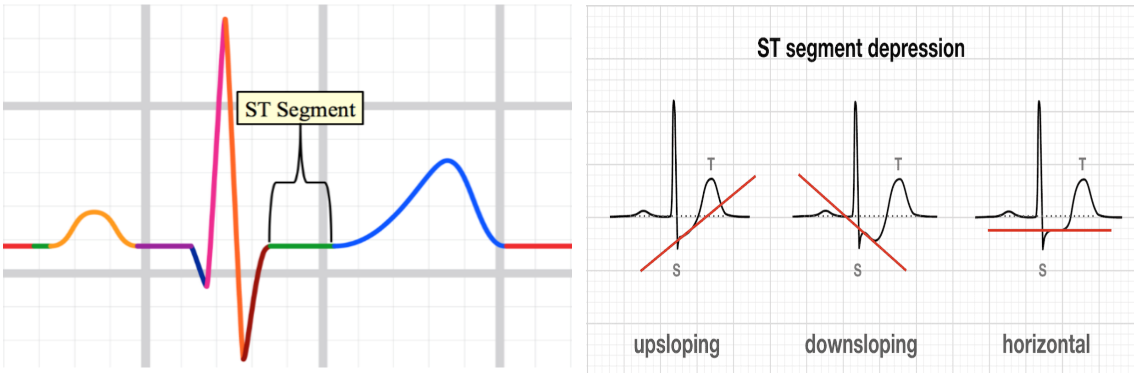

Figure (1) The ECG waveform and ST-segments (Mamun & Alouani, 2020)

Al-Ghary Journal of Economic and Administrative Sciences Vol. 22 (No.2) 2026 PP. 135- 150

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (22) عدد (2) 2026

In machine learning, classification prediction refers to the process of assigning categorical labels to unlabeled data instances based on learned patterns from training samples. The classifier's output represents a discrete class membership prediction for each input sample within a predefined label space.(Sekeroglu et al., 2022). In this article used five ML algorithm are:

2-1 Machine Learning (ML)

Machine learning (ML) is a subset of artificial intelligence that involves the use of algorithms and statistical models to enable computer systems to learn from data without being explicitly programmed. This technology has revolutionized various industries, including healthcare, finance, and transportation.The key advantage of ML is ability to process large amounts of data quickly and accurately, organizations to make data-driven decisions and improve their operations. ML has demonstrated its powerful predictive capabilities and parallel processing able for handling large numbers of variables. Furthermore, ML has derived variable screening mechanisms that can detect and interpret complex relationships between variables. Although most ML studies included the comparative analysis in determining the optimal model. (Qin et al., 2022; Sekeroglu et al., 2022).

The application of ML to electrocardiogram (ECG) analysis has significantly evolved Early foundational work demonstrated the immense potential of ML models applied to raw ECG signals. Previous studies have well demonstrated that ML may be employed as an effective research method for predicting. ML enhances cardiovascular disease (CVD) diagnosis and prognosis by detecting subtle patterns in ECG data beyond human capability. For STEMI diagnosis, ML models outperform traditional methods (Mamun & Alouani,

DOI: https://doi.org/10.36325/ghjec.v22i2.21032

2020; Mehrang et al., 2018), with interpretable approaches addressing the "black box" problem (Bodini et al., 2020; Jurado et al., 2022). In predictive modeling, ML stratifies risk for mortality (Kasim et al., 2024), post-PCI events (Chen et al., 2025), and heart failure subtypes using ECG indicators (Liu et al., 2025) and circadian features (Al Younis et al., 2024). Methodological studies compare algorithms (Nti et al., 2021; Sekeroglu et al., 2022) and enable population-level diagnostics (Kalmady et al., 2024) and data visualization (Nayak et al., 2021).

2-2 Adaptive Boost Algorithm (AdaBoost)

Adaptive Boosting (AdaBoost), introduced by Freund and Schapire (1996), is widely regarded as the first practical boosting algorithm for ensemble ML (Nti et al., 2021). This iterative method is computationally efficient, straightforward to implement, and requires minimal hyper parameter tuning specifically, only the number of iterations. Notably resistant to overfitting, AdaBoost excels at detecting outliers, particularly instances that are misclassified or inherently difficult to classify. The algorithm operates by sequentially generating weak classifiers and combining them through a weighted majority vote to construct a robust final classifier.

2-3 Gradient Boosting Algorithm (GB)

Gradient Boosting (GB) combines predictions from multiple decision trees to generate robust final outputs, making it one of the most powerful machine learning algorithms for both regression and classification tasks. As opposite standalone models, GB employs an ensemble approach in which decision trees are constructed sequentially, with each subsequent tree learning from and correcting the errors of its predecessors (Nti et al., 2021). The algorithm optimizes performance by iteratively minimizing the loss function of the weak learners (decision trees). This is achieved through gradient descent: at each step, the loss is computed, and a new tree is added or modified to reduce the residual errors, thereby progressively enhancing predictive accuracy (Sekeroglu et al., 2022).

2-4 Extreme Gradient Boosting (XGBoost)

The XGBoost (Extreme Gradient Boosting) algorithm, introduced by Chen and Guestrin,(Qin et al., 2022) is an advanced boosting technique designed to enhance model performance while mitigating overfitting through a regularization parameter. Like traditional Gradient Boosting (GB), XGBoost employs an ensemble tree approach, iteratively improving weak learners via gradient descent. However, XGBoost incorporates key optimizations to boost computational efficiency, reduce resource consumption, and deliver superior predictive accuracy. It is widely applicable to both regression and classification tasks in machine learning (Nti et al., 2021; Sekeroglu et al., 2022).

As an ensemble learning method, XGBoost constructs a robust predictive model by sequentially integrating multiple weak models (commonly applied in medical diagnostics, such as heart disease prediction). Each subsequent decision tree in XGBoost corrects errors from the previous one, refining the model iteratively. Additionally, XGBoost introduces several improvements over conventional gradient boosting, including enhanced speed, reduced overfitting, and better generalization capabilities (Al Younis et al., 2024).

2-5 Light Gradient Boosting Algorithm (LightGBM)

LightGBM achieves faster training speeds and greater computational efficiency compared to traditional gradient boosting methods (Nti et al., 2021). As a decision tree based gradient boosting framework, LightGBM

مجلة الغري للعلوم الاقتصادية والادارية
مجلد (22) عدد (2) 2026

enhances model performance while significantly reducing memory consumption. Which are widely used in gradient boosting decision trees. These techniques optimize data handling and feature processing, leading to improved scalability and reduced computational overhead (Bhoite et al., 2023).

2-6 CAT Boost Algorithm

CatBoost developed by Yandex demonstrates superior predictive accuracy compared to XGBoost and LightGBM, despite XGBoost's widespread industry adoption and LightGBM's notable computational efficiency advantages (Sekeroglu et al., 2022). As a member of the boosting family, CatBoost distinguishes itself through two key characteristics: (1) reduced dependency on extensive hyperparameter tuning, and (2) enhanced resistance to overfitting, which improves model generalizability. In terms of performance, CatBoost remains competitive with state-of-the-art machine learning algorithms.

3- Results

ML has established itself as a transformative paradigm in clinical medicine, significantly advancing diagnostic capabilities beyond traditional approaches. These algorithms demonstrate particular utility in medical domains through their ability to identify complex patterns within high-dimensional clinical data. This study utilized a retrospective cohort of (1000) patients with ST-segment abnormalities from Shar Hospital in Sulaymaniyah City. The dataset comprised comprehensive demographic and clinical variables, including:

Table (1) association between ST-segment, different demographic and clinical variables

Al-Ghary Journal of Economic and Administrative Sciences Vol. 22 (No.2) 2026 PP. 139- 150
مجلة الغري للعلوم الاقتصادية والادارية
مجلد (22) عدد (2) 2026

Variables	Classes	ST-depression f_i	ST-elevations f_i	p value
Age	20-30	0	23	0.004
	31-40	7	18	
	41-50	12	106	
	51-60	51	190	
	61-70	60	167	
	71-80	55	166	
	81-90	21	82	
	91-100	9	33	
Gender	Female	95	303	0.138
	Male	120	482	
Smoking	No	16	70	0.494
	Yes	199	715	
Hypertension	No	65	179	0.025
	Yes	150	606	
Chronic Cardiovascular	No	54	133	0.006
	Yes	161	652	
Diabetic	No	35	155	0.251
	Yes	180	630	
Dyslipidemia	No	65	179	0.025
	Yes	150	606	
Hypothyroidism	No	76	257	0.472
	Yes	139	528	

Not: f_i: frequency of each class, P value: by using Chi- square test.

The table above show the association between ST-segment depression and ST-segment elevations across different demographic and clinical variables, including age, gender, smoking status, hypertension, chronic cardiovascular disease (CVD), diabetes, dyslipidemia, and hypothyroidism. The data presented in Table (1) contain (1000) abnormality ST-segment collected from Shar Hospital in Sulaymaniyah City's, statistically significant relationships ($p < 0.05$) that may inform clinical risk stratification.

Age and ST-segment changes the distribution of ST-elevations was highest in patients aged class (51-60, f_i=190) and (61-70, f_i =167), while ST-depression was most prevalent in the (61-70, f_i =60) and (51-60, n=51) age groups. A significant p-value (p =0.004) suggests a strong association between age and ST-segment abnormalities, indicating that older adults are more likely to exhibit ischemic ECG changes.

Gender differences males exhibited higher frequencies of both ST-depression (f_i=120) and ST-elevations (f_i =482) compared to females (ST-depression: f_i =95; ST-elevations: n=303).However, the p-value (p = 0.138) was not statistically significant, suggesting that gender may not be a decisive factor in ST-segment changes.

Smoking Status a majority of patients presenting with ST-segment abnormalities were smokers (ST-depression: f_i =199; ST-elevations: f_i =715), but the p-value (p = 0.494) did not reach statistical significance. This implies that while smoking is common among these patients, it may not independently predict ST-segment changes.

Hypertension and Dyslipidemia both hypertension (p=0.025) and dyslipidemia (p=0.025) showed significant associations with ST-segment deviations. Patients with hypertension had notably higher occurrences of ST-elevations (f_i =606) compared to non-hypertensive patients (f_i =179).

Chronic Cardiovascular Disease (CVD) a strongly significant relationship (p=0.006) was observed between chronic CVD and ST-segment abnormalities. Patients with CVD had substantially more ST-elevations (f_i =652) and ST-depressions (f_i =161) than those without CVD.

Diabetes and Hypothyroidism Diabetes did not show a significant association (p=0.251), despite a higher raw count of abnormalities in diabetic patients. Similarly, hypothyroidism had no significant impact (p=0.472) on ST-segment changes.

DOI: https://doi.org/10.36325/ghjec.v22i2.21032

Al-Ghary Journal of Economic and Administrative Sciences Vol. 22 (No.2) 2026 PP. 141- 150

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (22) عدد (2) 2026

We conducted a comprehensive evaluation of five boosting algorithms (XGBoost, CatBoost, LightGBM, AdaBoost, and Gradient Boosting) using four distinct training-testing split ratios (70-30, 75-25, 80-20, and 85-15). Model performance was assessed across four key classification metrics: accuracy, precision, recall, and F1-score. As shown in Figure (2) , the comparative analysis of F1-scores - the harmonic mean of precision and recall reveals distinct performance patterns across split configurations, with AdaBoost demonstrating particularly robust performance (F1 = 0.880) while maintaining perfect recall (1.0000) across all data partitions.

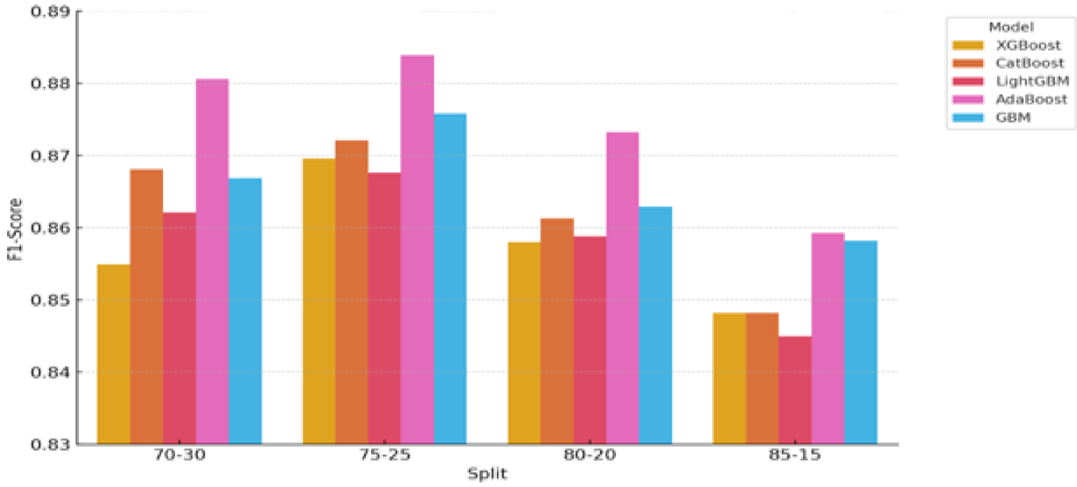

Figure (2) compares F1-Scores of five models to four different train-test splits

AdaBoost consistently achieves the highest F1-scores across all splits (at 0.8839 in the 75-25 split), underscoring its superiority in balancing precision and recall. GB and CatBoost demonstrate competitive performance, with F1-scores closely trailing AdaBoost (GB: 0.8758, CatBoost: 0.8721 in the 75-25 split). Boost and LightGB exhibit marginally lower but stable performance, with LightGB showing slight advantages in certain splits (e.g., 80-20).

مجلة الغري للعلوم الاقتصادية والادارية
مجلد (22) عدد (2) 2026

The (75-25) split emerges as the optimal configuration for most models, likely due to its balance between sufficient training data and reliable evaluation. AdaBoost's F1-score peaks at this split (0.8839), declining slightly to [0.8806 (70-30) and 0.8732 (80-20)]. Similar trends are observed for GB and CatBoost, suggesting this split mitigates overfitting while preserving generalizability. The (85-15) split shows performance degradation for all models (e.g., AdaBoost drops to 0.8593).

Table 2. Comparative performance evaluation of five ML models and train-test split

Performance	train-test	Models				
		XGBoost	CatBoost	LightGB	AdaBoost	GB
Precision	70-30	0.7865	**0.7909**	0.7867	0.7867	0.7862
	75-25	0.7950	**0.7958**	0.7917	0.7920	0.7918
	80-20	0.7789	**0.7801**	0.7760	0.7750	0.7744
	85-15	**0.7569**	**0.7569**	0.7517	0.7533	0.7568
Recall	70-30	0.9364	0.9619	0.9534	**1.0000**	0.9661
	75-25	0.9596	0.9646	0.9596	**1.0000**	0.9798
	80-20	0.9548	0.9613	0.9613	**1.0000**	0.9742
	85-15	0.9646	0.9646	0.9646	**1.0000**	0.9912
F1-Score	70-30	0.8549	0.8681	0.8621	**0.8806**	0.8669
	75-25	0.8696	0.8721	0.8676	**0.8839**	0.8758
	80-20	0.8580	0.8613	0.8588	**0.8732**	0.8629
	85-15	0.8482	0.8482	0.8450	**0.8593**	0.8582

Al-Ghary Journal of Economic and Administrative Sciences Vol. 22 (No.2) 2026 PP. 143- 150

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (22) عدد (2) 2026

Performance	train-test	Models				
		XGBoost	CatBoost	LightGB	AdaBoost	GB
Accuracy	70-30	0.7500	0.7700	0.7600	**0.7867**	0.7667
	75-25	0.7720	0.7760	0.7680	**0.7920**	0.7800
	80-20	0.7550	0.7600	0.7550	**0.7750**	0.7600
	85-15	0.7400	0.7400	0.7333	**0.7533**	**0.7533**

At Training 70 - Testing 30:

AdaBoost achieved the highest accuracy (0.7867), F1-Score (0.8806) and perfect recall (1.0000), indicating strong sensitivity to positive instances. CatBoost and GB followed closely in F1-Score (0.8681 and 0.8669, respectively), balancing precision and recall effectively.

Training 75 - Testing 25:

AdaBoost achieves the highest accuracy (0.7920) and perfect recall (1.0000), reinforcing its robustness in classification. CatBoost demonstrated the best precision (0.7958) and a competitive F1-Score (0.8721), GB also performs well with high recall (0.9798) and F1-Score (0.8758).

Training 80 - Testing 20:

AdaBoost remained dominant in recall (1.0000) and F1-Score (0.8732), though with slightly lower precision (0.7750). CatBoost and GB maintained stable performance, with CatBoost achieving a marginally better F1-Score (0.8613).

Training 85 - Testing 15:

AdaBoost continued its trend of perfect recall, but with reduced precision (0.7533), leading to an F1-Score of (0.8593). XGBoost and CatBoost showed

DOI: https://doi.org/10.36325/ghjec.v22i2.21032

Al-Ghary Journal of Economic and Administrative Sciences Vol. 22 (No.2) 2026 PP. 144- 150

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (22) عدد (2) 2026

similar performance in precision and recall, though overall metrics declined slightly with less training data.

Across all splits, AdaBoost consistently achieved the highest recall and F1-Score, indicating superior sensitivity to positive cases. However, CatBoost and GB provided better precision-recall trade-offs, making them suitable for applications where false positives must be minimized.

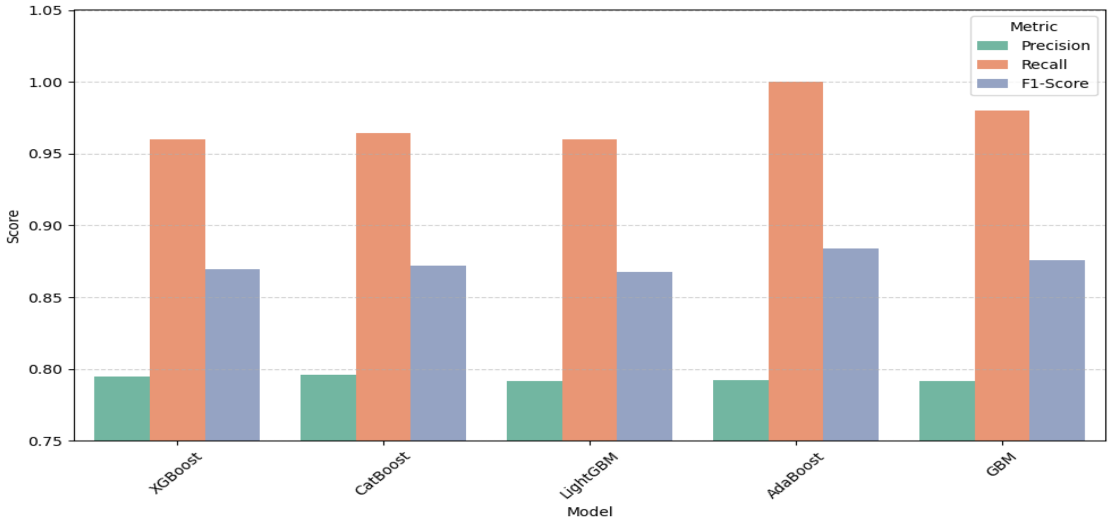

Figure (3) comparative performance for models

The Figure above presents a comparative performance analysis of five ensemble learning models (XGBoost, CatBoost, LightGBM, AdaBoost, and Gradient Boosting) across four key classification metrics: precision, recall, and F1-score but do not show changes in accuracy, as the accuracy values for the five different models are approximately equal (around 0.7). The visualization reveals distinct performance characteristics among the evaluated algorithms, with AdaBoost demonstrating particular strengths in recall while maintaining competitive performance across other metrics. Characteristics that inform model selection for clinical prediction tasks. Best model (F1-Score): AdaBoost is the top performer with an F1-score of

(0.8839) and perfect recall (1.0000). GB and CatBoost also perform very well with high precision and recall.

Best Training-Testing Split: (75 Training - 25 Testing) is the optimal split for this dataset, as it provides: Highest accuracy (AdaBoost: 0.7920). Best balance between precision and recall (CatBoost: 0.7958 Precision, 0.9646 Recall). Strongest overall F1-Scores. If maximizing recall is the priority, AdaBoost with (75-25) split is the best choice. If balancing precision and recall is more important, CatBoost or GBM with (75-25) split is preferable. AdaBoost achieves perfect recall (1.0000) in all splits and the highest F1-scores. GB shows the highest recall (0.9912) in the (85-15) split. (75-25) split yields the best balance of performance across all models. Precision remains relatively stable (0.75–0.79) for all models. Best Precision: CatBoost (75–25), best recall: AdaBoost (all splits with 1.0000), best F1-Score, Accuracy: AdaBoost (especially at 75–25 split).

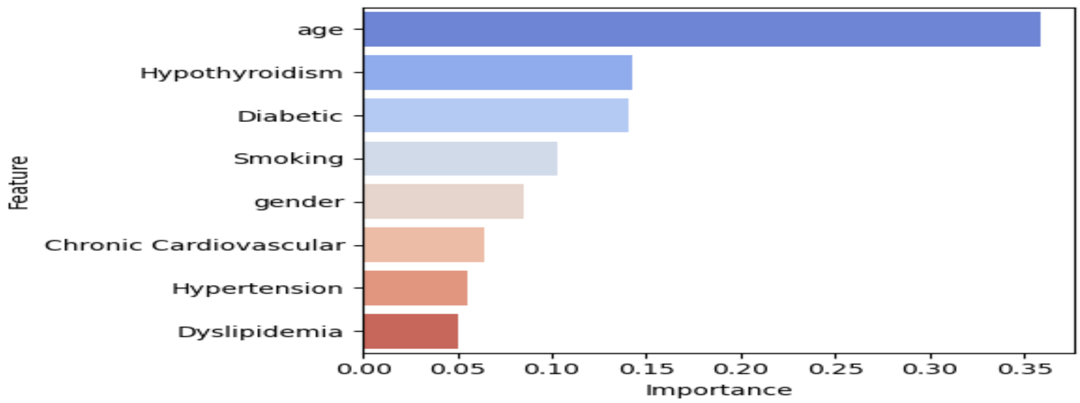

Figure (4) feature importance

Figure above illustrates the feature importance scores derived from the AdaBoost classifier, an ensemble learning algorithm that sequentially constructs a robust predictive model by aggregating multiple weak learners. Among the evaluated features, age emerged as the most influential predictor, contributing approximately 40% to the overall model, which is substantially

DOI: https://doi.org/10.36325/ghjec.v22i2.21032

higher than any other variable. This suggests that age plays a central role in predicting the target outcome within the dataset, potentially reflecting age-related risk escalation for the condition under study.

Following age, the next most important features include Hypothyroidism**,** Diabetic status, and Smoking**,** each providing meaningful contributions to the model. These features are known risk factors in many clinical domains and may be strongly associated with the outcome variable in the dataset. Their moderate importance suggests they help the model refine predictions in conjunction with age.

In contrast, features such as Hypertension**,** Dyslipidemia**,** Chronic Cardiovascular conditions**,** and gender exhibited relatively lower importance scores. While these variables may still hold clinical significance, their lower contribution in the model indicates that they may not provide as much discriminative power for prediction in the presence of more dominant variables like age.

Table (3) average performance and different training- testing split

Model	Avg Accuracy	Avg Precision	Avg Recall	Avg F1-Score
XGBoost	~0.762	~0.788	~0.952	~0.863
CatBoost	~0.768	**~0.791**	~0.963	~0.868
LightGBM	~0.758	~0.786	~0.958	~0.864
AdaBoost	**~0.786**	~0.786	**1.000**	**~0.880**
GBM	~0.772	~0.787	~0.972	~0.870

In table (3) best overall model AdaBoost Highest F1-Score (~0.880) and Recall (1.000) in all tests. That means no false negatives: it never misses a positive case, which is excellent in medical diagnosis, fraud detection, etc. Precision is slightly lower, meaning it has some false positives (but not many). Second best: GB: Very close to AdaBoost in F1-Score (~0.87) and Recall (~0.972). Slightly better Precision than AdaBoost so it balances better

مجلة الغري للعلوم الاقتصادية والادارية
مجلد (22) عدد (2) 2026

if false positives matter too. CatBoost slightly edges out XGBoost. Strong, but F1-Scores and Recall are consistently lower than AdaBoost and GB. LightGB, Still good, but slightly lower across all metrics. Fast training and lower resource use make it a great option for very large datasets, but performance is just behind others here.

4- Conclusions

The findings suggest that age, hypertension, dyslipidemia, and chronic CVD are significant predictors of ST-segment deviations, aligning with existing literature on cardiovascular risk factors. The lack of significance in smoking, diabetes, and gender may indicate that these factors require further stratification.This study evaluated five boosting algorithms under varying training-testing splits to assess their classification performance. The findings are: best splits, (75–25) and AdaBoost demonstrated the highest recall and F1-Score in all scenarios, CatBoost and GB offered a more balanced performance, particularly in precision, XGBoost and LightGB showed competitive but slightly lower results, with performance declining.

Overall best model the performance hierarchy (AdaBoost > GB > CatBoost > XGBoost > LightGB) suggests that adaptive boosting methods may be particularly well suited for clinical prediction tasks. Feature importance results reinforce the centrality of age as the most influential variable, aligning with clinical understanding that age is a major risk factor in many health related conditions. The identification of Hypothyroidism, Diabetes, and Smoking as key contributing factors further supports their relevance in the prediction of the outcome of interest.

Disclosure Statement: This research constitutes a component of an ongoing PhD dissertation in the Department of Statistics and Informatics at the

مجلة الغري للعلوم الاقتصادية والادارية
مجلد (22) عدد (2) 2026

College of Administration and Economics, University of Sulaimani, Sulaimaniyah, Iraq.

5- Reference

1- Al Younis, S. M., Hadjileontiadis, L. J., Khandoker, A. H., Stefanini, C., Soulaidopoulos, S., Arsenos, P., Doundoulakis, I., Gatzoulis, K. A., & Tsioufis, K. (2024). Prediction of heart failure patients with distinct left ventricular ejection fraction levels using circadian ECG features and machine learning. *PLoS ONE*, *19*(5 May), 8–9. https://doi.org/10.1371/journal.pone.0302639

2- Bhoite, S. N., Gadekar, V. D., Kapadnis, S. V., & Ghuge, P. R. (2023). Churn Prediction using Machine Learning Models. *International Journal for Research in Applied Science and Engineering Technology*, *11*(5), 2233–2236. https://doi.org/10.22214/ijraset.2023.52101

3- Bodini, M., Rivolta, M. W., & Sassi, R. (2020). Interpretability Analysis of Machine Learning Algorithms in the Detection of ST-Elevation Myocardial Infarction. *Computing in Cardiology*, *2020-Septe*, 1–4. https://doi.org/10.22489/CinC.2020.403

4- Chen, M., Sun, C., Yang, L., Zhang, T., Zhang, J., & Chen, C. (2025). Application of Machine Learning Algorithms in Predicting Major Adverse Cardiovascular Events after Percutaneous Coronary Intervention in Patients with New-Onset ST-Segment Elevation Myocardial Infarction. *Reviews in Cardiovascular Medicine*, *26*(2). https://doi.org/10.31083/RCM25758

5- Jurado, I. C., Fedjajevs, A., Vanschoren, J., & Brombacher, A. (2022). Interpretable Assessment of ST-Segment Deviation in ECG Time Series. *Sensors*, *22*(13), 1–24. https://doi.org/10.3390/s22134919

6- Kalmady, S. V., Salimi, A., Sun, W., Sepehrvand, N., Nademi, Y., Bainey, K., Ezekowitz, J., Hindle, A., McAlister, F., Greiner, R., Sandhu, R., & Kaul, P. (2024). Development and validation of machine learning algorithms based on electrocardiograms for cardiovascular diagnoses at the population level. *Npj Digital Medicine*, *7*(1). https://doi.org/10.1038/s41746-024-01130-8

7- Kasim, S., Amir Rudin, P. N. F., Malek, S., Ibrahim, K. S., Wan Ahmad, W. A., Fong, A. Y. Y., Lin, W. Y., Aziz, F., & Ibrahim, N. (2024). Ensemble machine learning for predicting in-hospital mortality in Asian women with ST-elevation myocardial infarction (STEMI). *Scientific Reports*, *14*(1), 1–16. https://doi.org/10.1038/s41598-024-61151-x

8- Liu, J., Zhu, D., Deng, L., & Chen, X. (2025). Predictive Modeling of Heart Failure Outcomes Using ECG Monitoring Indicators and Machine Learning. *Annals of Noninvasive Electrocardiology*, *30*(4), 1–7. https://doi.org/10.1111/anec.70097

9- Mamun, M. M. R. K., & Alouani, A. (2020). Diagnosis of STEMI and Non-STEMI Heart Attack using Nature-inspired Swarm Intelligence and Deep Learning Techniques. *Journal of Biomedical Engineering and Biosciences*, *7*. https://doi.org/10.11159/jbeb.2020.001

10- Mehrang, S., Jafari Tadi, M., Kaisti, M., Lahdenoja, O., Vasankari, T., Kiviniemi, T., Airaksinen, J., Koivisto, T., & Pankaala, M. (2018). Machine Learning Based Classification of Myocardial Infarction Conditions Using Smartphone-Derived Seismo- and Gyrocardiography. *Computing in Cardiology*, *2018-Septe*(Mi), 1–4. https://doi.org/10.22489/CinC.2018.110

11- Nayak, K. R., Skupin, A., Schempp, T., Garberich, R., Bhavnani, S. P., & Henry, T. (2021). Machine learning for holistic visualization of STEMI registry data. *Journal of Biomedical Informatics*, *121*(July), 1–11. https://doi.org/10.1016/j.jbi.2021.103869

12- Nti, I. K., Aning, J., Ayawli, B. B. K., Frimpong, K., Appiah, A. Y., & ... (2021). A Comparative Empirical Analysis of 21 Machine Learning Algorithms for Real-World Applications in Diverse Domains. *Journal of Big Dats*. https://www.researchsquare.com/article/rs-518365/latest.pdf

13- Qin, Y., Wu, J., Xiao, W., Wang, K., Huang, A., Liu, B., Yu, J., Li, C., Yu, F., & Ren, Z. (2022). Machine Learning Models for Data-Driven Prediction of Diabetes by Lifestyle Type. *International Journal of Environmental Research and Public Health*, *19*(22). https://doi.org/10.3390/ijerph192215027

14- Sekeroglu, B., Ever, Y. K., Dimililer, K., & Al-Turjman, F. (2022). Comparative Evaluation and Comprehensive Analysis of Machine Learning Models for Regression Problems. *Data Intelligence*, *4*(3), 620–652. https://doi.org/10.1162/dint_a_00155

15- Symposium, I. (2024). *or ap er Sy mp os or g P ap er of Sy mp os*. https://doi.org/10.2507/33nd.daaam.proceedings.xxx

16- Vogel, B., Claessen, B. E., Arnold, S. V., Chan, D., Cohen, D. J., Giannitsis, E., Gibson, C. M., Goto, S., Katus, H. A., Kerneis, M., Kimura, T., Kunadian, V., Pinto, D. S., Shiomi, H., Spertus, J. A., Steg, P. G., & Mehran, R. (2019). ST-segment elevation myocardial infarction. *Nature Reviews Disease Primers*, *5*(1), 1–20. https://doi.org/10.1038/s41572-019-0090-3