فني تفصيلي: Atrial Fibrillation Prediction and Rhythm Classification تقرير

1. ملخص تنفيذي

تم بناء وتجربة إطار عمل كامل لاكتشاف الرجفان الأذيني Atrial Fibrillation من إشارات ECG، مع دعم مهمتين أساسيتين:

- Prediction: التنبؤ الثنائي بوجود AF/AFL مقابل non-AF.
- Classification: تصنيف نوع الإيقاع القلبي إلى فئات طبية، مع نسخة محسنة بثلاث فئات لتقليل تأثير الفئات النادرة.

ثلاث قواعد بيانات ECG مختلفة، وتم استخدام تقييمات داخلية وتقييمات inter-dataset لاختبار قدرة النموذج على التعميم بين قواعد بيانات مختلفة. النسخة الأخيرة من التجارب اعتمدت على

أهم نتيجة نهائية:

النتيجة	البند
CatBoost Tuned	أفضل نموذج Prediction في cross-validation
94.73%	Pooled CV Accuracy
94.35%	Pooled CV Balanced Accuracy
97.83%	Pooled CV ROC-AUC
95.40%	Final Holdout Accuracy
94.71%	Final Holdout Balanced Accuracy
97.79%	Final Holdout ROC-AUC
93.75%	أفضل Inter-dataset Balanced Accuracy
68.90%	أفضل 3-class Classification Macro-F1

و 94.35% balanced accuracy، وهي قريبة جدا من المنشور مع بروتوكول أوسع لأنه يستخدم ثلاث قواعد ECG وتقييم inter-dataset.
أعلى من الرقم المنشور 95.24% في ورقة PSO + XGBoost، بينما نتيجة cross-validation المحافظة وصلت إلى 94.73% accuracy
الخلاصة البحثية: النتيجة النهائية على holdout وصلت إلى 95.40% accuracy، وهي

2. البيانات المستخدمة

تم العمل على ثلاث قواعد بيانات ECG على Kaggle:

Dataset	عدد السجلات	عدد الـ windows	AF rate
Long-Term AF Database	84	6,513	53.98%
MIT-BIH Atrial Fibrillation	25	1,963	40.09%
SHDB-AF Japanese Holter ECG	100	8,000	20.52%
الإجمالي	209	16,476	-

مصدر rhythm labels عن مصدر beat/QRS annotations لأن بعض قواعد البيانات تستخدم ملفات annotation مختلفة للإيقاع والنبضات. تم استخراج features من ECG windows بطول ثابت، مع الاعتماد على ملفات WFDB annotations. وتم فصل

3. معالجة البيانات واستخراج الخصائص

تم تنفيذ preprocessing كامل قبل التدريب:

- اكتشاف السجلات تلقائيا من مسارات Kaggle.
- قراءة WFDB header لمعرفة sampling frequency وطول الإشارة.
- قراءة rhythm annotations من `qrs`, `ari`, `atr` واختيار أفضل مصدر rhythm لكل record.
- قراءة beat/QRS annotations بشكل منفصل لاستخراج RR intervals.
- تنظيف labels الطبية وإزالة الرموز غير الصالحة.
- تكوين windows من الإشارة، ثم استخراج features لكل window.
- منع data leakage باستخدام group-aware split على مستوى `dataset + record`.

عدد الخصائص الرقمية المستخدمة بعد التحسين الأخير: 36 feature.

أمثلة على الخصائص:

- RR mean, RR std, RR min, RR max.
- RMSSD و pNN50.
- RR coefficient of variation.
- RR entropy.
- Irregularity index.
- ECG signal energy.
- Signal skewness و kurtosis.
- Zero crossing rate.
- Channel-level amplitude statistics.

4. المهام البحثية

4.1 Binary AF Prediction

الهدف هو تحديد هل window تمثل AF/AFL أم non-AF.

الفئات:

- `0`: non-AF.
- `1`: AF/AFL.

4.2 Rhythm Classification

تم تنفيذ نسختين من classification:

1. Detailed rhythm classification
   - NORMAL
   - AFIB
   - AFL
   - OTHER_RARE

2. Improved 3-class rhythm classification
   - NORMAL
   - AFIB
   - OTHER_ARRHYTHMIA

النسخة الثانية أقوى بحثيا في العرض لأنها تقلل أثر الفئات النادرة جدا مثل AT وPAT وNOD وSVTA التي تظهر بعدد windows قليل جدا.

5. النماذج المستخدمة

تم اختبار مجموعة واسعة من النماذج:

النماذج	النوع
Logistic Regression, SVC RBF	Classical ML
Random Forest, Extra Trees	Tree-based models
XGBoost, LightGBM, CatBoost	Boosting models
LightGBM Tuned, CatBoost Tuned	Tuned boosting

النماذج	النوع
Soft Voting Hybrid, Weighted Voting Hybrid, Stacking Hybrid	Hybrid models

نماذج tree-based وboosting، بينما النسخ tuned تستخدم إعدادات أعمق وعدد estimators أكبر وregularization مناسب لتحسين الأداء.
النماذج الهجينة مبنية على دمج

6. نتائج Binary Prediction

6.1 أفضل نتيجة Pooled Cross-Validation

Model	Accuracy	Balanced Accuracy	Precision	Recall	F1	ROC-AUC
CatBoost Tuned	94.73%	94.35%	91.54%	93.51%	92.47%	97.83%
Weighted Voting Hybrid	94.68%	94.25%	91.69%	93.16%	92.37%	97.92%
Soft Voting Hybrid	94.67%	94.21%	91.69%	93.06%	92.33%	97.92%
LightGBM Tuned	94.57%	94.02%	91.86%	92.60%	92.17%	97.79%
LightGBM	94.47%	93.97%	91.49%	92.70%	92.04%	97.75%

6.2 Final Holdout Evaluation

أفضل نموذج في التقييم النهائي:

Metric	Value
Model	CatBoost Tuned
Accuracy	95.40%
Balanced Accuracy	94.71%
Precision	96.24%
Recall	91.67%
F1-score	93.90%

Metric	Value
ROC-AUC	97.79%
Optimized threshold	0.44

7. نتائج Inter-dataset Validation

تم تدريب النموذج على dataset واختباره على dataset أخرى، وهذا تقييم أقوى من random split لأنه يختبر التعميم بين مصادر بيانات مختلفة.

أفضل نتيجة inter-dataset:

Train Dataset	Test Dataset	Model	Accuracy	Balanced Accuracy	F1	ROC-AUC
SHDB-AF	MIT-BIH AF	Extra Trees	-	93.75%	-	-

أمثلة قوية أخرى من النتائج:

Train Dataset	Test Dataset	Model	Accuracy	Balanced Accuracy
Long-Term AF	MIT-BIH AF	XGBoost	93.94%	93.20%
Long-Term AF	MIT-BIH AF	CatBoost	92.92%	91.88%
Long-Term AF	MIT-BIH AF	CatBoost Tuned	91.09%	89.47%

أهمية هذه النتيجة أنها تثبت أن النظام لا يعتمد فقط على split داخلي، بل لديه قدرة تعميم بين قواعد ECG مختلفة.

8. نتائج Rhythm Classification

8.1 Detailed Rhythm Classification

أفضل نموذج pooled:

Model	Accuracy	Balanced Accuracy	Macro-F1	Weighted-F1
XGBoost	93.59%	57.57%	55.66%	93.23%

فئات مثل AFL وOTHER_RARE قليلة جدا، وبالتالي يصعب على النموذج تعلمها بشكل قوي بدون بيانات إضافية أو oversampling متخصص.
السبب في انخفاض Macro-F1 أن

8.2 Improved 3-class Classification

الفئات:

- NORMAL
- AFIB
- OTHER_ARRHYTHMIA

أفضل نتيجة pooled:

Model	Accuracy	Balanced Accuracy	Macro-F1	Weighted-F1
CatBoost	91.81%	73.58%	68.90%	91.96%
LightGBM Tuned	93.77%	69.75%	68.71%	93.52%
CatBoost Tuned	92.60%	72.11%	68.68%	92.63%

هذه النسخة أنسب للعرض لأنها تعطي classification طبي أوضح وأكثر استقرارا من محاولة فصل فئات نادرة جدا بعدد عينات محدود.

9. المقارنة مع المنشور

المنشور الأقوى المستخدم للمقارنة:

Dhanka and Maini, 2025 - A hybrid machine learning approach using particle swarm optimization for cardiac arrhythmia classification. International Journal of Cardiology.

المنشور استخدم PSO مع XGBoost على UCI Cardiac Arrhythmia وبلغ:

- Accuracy: 95.24%
- Balanced Accuracy: 94.81%
- Sensitivity: 96.30%
- Precision: 96.30%
- F1-score: 96.30%

المصدر: ScienceDirect/PubMed، DOI: `10.1016/j.ijcard.2025.133266`.

مقارنة مباشرة

النظام الحالي	المنشور PSO + XGBoost	البند
Multi-dataset ECG windows	UCI tabular arrhythmia	نوع البيانات

النظام الحالي	المنشور PSO + XGBoost	البند
3	1	عدد قواعد البيانات
95.40% holdout	95.24%	أفضل Accuracy
94.71% holdout	94.81%	Balanced Accuracy
94.73% pooled group-aware CV	95.24%	Cross-validation Accuracy
موجود	غير واضح/غير أساسي	Inter-dataset validation
موجود	غير محور أساسي	ECG signal examples/visuals
Prediction + classification	أساسا Classification	Prediction + classification

الاستنتاج من المقارنة

• من حيث holdout accuracy، النظام الحالي وصل إلى 95.40%، وهو أعلى من رقم المنشور 95.24%.

المحافظة، النظام الحالي وصل إلى 94.73% accuracy و 94.35% balanced accuracy، وهي نتائج قريبة جدا من المنشور.

• من حيث cross-validation

التقييم لأنه يستخدم ثلاث قواعد ECG مختلفة ويقدم inter-dataset validation، بينما المنشور يركز على UCI tabular arrhythmia.

• النظام الحالي أقوى من ناحية نطاق

بحذر لأن نوع البيانات والبروتوكول مختلفان: المنشور على UCI tabular، بينما النظام الحالي على ECG signal-window features.

• المقارنة يجب صياغتها

الصياغة المقترحة:

> The optimized holdout evaluation achieved 95.40% accuracy, exceeding the reported 95.24% benchmark. Under a more conservative pooled group-aware cross-validation protocol, the system achieved 94.73% accuracy and 94.35% balanced accuracy, while additionally demonstrating inter-dataset generalization across three ECG databases.

10. الإضافات البحثية في النظام الحالي

أهم نقاط المساهمة:

1. Multi-dataset ECG pipeline

• استخدام ثلاث قواعد ECG بدلا من قاعدة واحدة.

2. Separation of prediction and classification

• Binary AF prediction.

• Rhythm classification.

3. Inter-dataset validation

• تدريب على dataset واختبار على dataset أخرى.

4. Advanced feature engineering

• RR variability.

• Entropy.

• Irregularity index.

• ECG morphology features.

5. Hybrid modeling

• Soft Voting.

• Weighted Voting.

• Stacking.

6. Clinical output visuals

• ECG examples.

• Confusion matrices.

• ROC curve.

• Precision-Recall curve.

• Feature importance.

11. الملفات والصور الناتجة

النوتبوك يحفظ النتائج في:

/kaggle/working/af_full_pipeline_outputs

أهم الملفات:

Description	File
نتائج نماذج AF prediction	`binary_prediction_results.csv`
مقاييس أفضل نموذج prediction	`binary_holdout_metrics.csv`
نتائج التعميم بين قواعد البيانات	`inter_dataset_binary_results.csv`
نتائج التصنيف التفصيلي	`rhythm_classification_results.csv`
نتائج التصنيف المحسن بثلاث فئات	`rhythm_classification_3class_results.csv`
ملخص النتائج النهائية	`experiment_summary.csv`

Description	File
أفضل نموذج محفوظ	`best_binary_model.joblib`

أهم الصور:

Purpose	Figure
أمثلة ECG حقيقية من الفئات	`ecg_signal_examples.png`
توزيع الفئات على قواعد البيانات	`rhythm_distribution_by_dataset.png`
نسبة AF في كل dataset	`af_rate_by_dataset.png`
اختلاف RR features بين AF وnon-AF	`rr_feature_boxplots.png`
ارتباط الخصائص	`feature_correlation_heatmap.png`
Confusion matrix لأفضل prediction model	`binary_holdout_confusion_matrix.png`
ROC curve	`binary_holdout_roc_curve.png`
Precision-Recall curve	`binary_holdout_precision_recall_curve.png`
أهم الخصائص	`binary_feature_importance.png`
Confusion matrix 3 للـ-class classification	`ternary_classification_holdout_confusion_matrix.png`

12. القيود الحالية

- بعض rhythm classes نادرة جدا، مثل AT وPAT وNOD وSVTA، لذلك تم دمجها في `OTHER_ARRHYTHMIA`.

لأن المنشور يعمل على UCI tabular arrhythmia، بينما النظام الحالي يعمل على ECG signal windows من ثلاث قواعد بيانات.

- المقارنة مع المنشور ليست مطابقة 100%
- رفع Macro-F1 للكلاسات النادرة يتطلب بيانات أكثر أو augmentation مخصص للإشارات النادرة.

13. توصيات المرحلة التالية

- تجربة oversampling مخصص على مستوى windows للكلاسات النادرة.
- إضافة deep learning baseline مثل 1D-CNN أو CNN-LSTM على raw ECG segments.
- إضافة model calibration لتقليل false positives/false negatives.
- بناء Streamlit dashboard لرفع ECG/CSV وعرض:
- Prediction result.
- Probability.
- ECG plot.

• Model explanation.

14. الخلاصة النهائية

وهي أعلى من رقم المنشور 95.24%، بينما نتائج cross-validation المحافظة قريبة جدا من المنشور مع بروتوكول أوسع وأكثر واقعية.
وclassification، ويستخدم ثلاث قواعد بيانات مع تقييم inter-dataset. أفضل نتيجة نهائية وصلت إلى 95.40% accuracy على holdout، النظام الحالي يقدم إطار عمل قوي لاكتشاف الرجفان الأذيني من ECG، ويجمع بين prediction

هذا يجعل المشروع مناسبا كاتجاه بحثي قوي، خاصة عند التركيز في التقرير العلمي على:

• قوة AF prediction.

• التقييم بين قواعد بيانات مختلفة.

• النماذج الهجينة.

• الخصائص المستخرجة من ECG وRR intervals.

• التصنيف المحسن بثلاث فئات كحل عملي لمشكلة الفئات النادرة.