



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** X **Month of publication:** October 2026

DOI: <https://doi.org/10.22214/ijraset.2026.85119>

www.ijraset.com

Call: ☎ 08813907089

E-mail ID: ijraset@gmail.com

HbA1c-Informed Pattern Recognition for Diabetes Mellitus Classification and Hospital Readmission Risk Analysis

Sarita B. Ghansawant¹, Abhijeet G. Dhepe², Dr. Mukta G. Dhopeswarkar³

Department of Computer Science and IT, Dr. Babasaheb Ambedkar Marathwada University, Chhatrapati Sambhajinagar

Abstract: Glycated hemoglobin (HbA1c) is an established indicator of longer-term glycaemic exposure, but hospital readmission risk in patients with diabetes depends on additional clinical and healthcare-utilization factors. This paper presents a reproducible pattern-recognition framework that incorporates HbA1c and structured clinical variables from the Diabetes 130-US Hospitals for Years 1999–2008 dataset. The resource contains 101,766 inpatient encounters across 130 U.S. hospitals and integrated delivery networks. The proposed analysis addresses missing values, categorical encoding, patient-level data leakage, class imbalance, and model calibration. We compare logistic regression, decision trees, random forests, support vector machines, and gradient-boosting classifiers. Because the supplied source paper does not contain reproducible experimental predictions or raw model outputs, no new accuracy or area-under-the-curve values are claimed. Instead, the study motivates a transparent model-validation protocol and uses published performance estimates strictly as contextual benchmarks. The framework is intended for research and clinician-supported risk stratification, not autonomous diagnosis.

Index Terms: HbA1c, diabetes mellitus, pattern recognition, machine learning, electronic health records, hospital readmission, clinical decision support.

I. INTRODUCTION

Diabetes mellitus places sustained demands on health services because glycaemic control, comorbid disease, medication management, and care transitions interact over time. HbA1c offers a clinically useful summary of glucose exposure over approximately two to three months, yet its interpretation requires awareness of conditions that affect erythrocyte survival or hemoglobin composition [1].

The Diabetes 130-US Hospitals dataset provides a multivariate setting for evaluating whether HbA1c status adds predictive information when combined with demographic characteristics, diagnostic burden, prior utilization, and medication-related variables [2]. This paper therefore distinguishes HbA1c-status classification from the primary task of predicting 30-day readmission. The first classifies observed glycaemic categories; the second evaluates risk following discharge.

The present contribution is a methods-oriented research framework, not an original trained-model report. It specifies an evaluation plan that emphasizes patient-separated validation, discrimination, calibration, and minority-class sensitivity, thereby reducing the risk of misleading performance claims.

II. RESEARCH OBJECTIVES AND CONTRIBUTIONS

The original project describes classification of “hemoglobin-dependent diabetes mellitus” using pattern-recognition techniques. Because this is not an established clinical diabetes subtype, the term is used here only to indicate analysis informed by hemoglobin glycation and HbA1c. The primary objective is to specify a reliable prediction workflow for early hospital readmission among people with diabetes, with HbA1c treated as one component of their clinical profile. A secondary task—prediction of recorded HbA1c categories from clinical characteristics—can be evaluated only in encounters with observed HbA1c measurements. The two endpoints are not interchangeable, and neither should be described as diagnosing a new disease category.

The contributions are (1) a clinically grounded interpretation of HbA1c and its limitations; (2) an explicit data dictionary and missing-data plan for the UCI resource; (3) patient-group-aware preprocessing and validation procedures designed to reduce information leakage; (4) a prespecified comparison of linear and nonlinear classifiers; and (5) a transparent evaluation plan incorporating discrimination, calibration, class imbalance, subgroup reliability, and actionable clinical utility.

III. CLINICAL AND COMPUTATIONAL BACKGROUND

A. HbA1c as an Integrated Glycaemic Marker

HbA1c arises from nonenzymatic glycation of hemoglobin and reflects cumulative glycaemic exposure over approximately two to three months, with disproportionate contribution from recent weeks. In nonpregnant individuals, current ADA criteria permit an A1c value of 6.5% or higher as one diagnostic route for diabetes, while 5.7%–6.4% is associated with prediabetes. Appropriate assay methods and confirmatory tests remain necessary where hyperglycemia is not unequivocal [1]. These thresholds are clinical criteria, rather than directly usable numerical cutoffs for a dataset that contains only grouped A1c results.

HbA1c may be misleading in conditions affecting erythrocyte survival or hemoglobin composition, including some anemias, hemoglobin variants, blood loss, and transfusion. Disagreement between HbA1c and glucose measurements warrants clinical assessment [1]. For this reason, missing or atypical A1c results must not be imputed as normal glycaemic status.

HbA1c: chronic glycaemic exposure

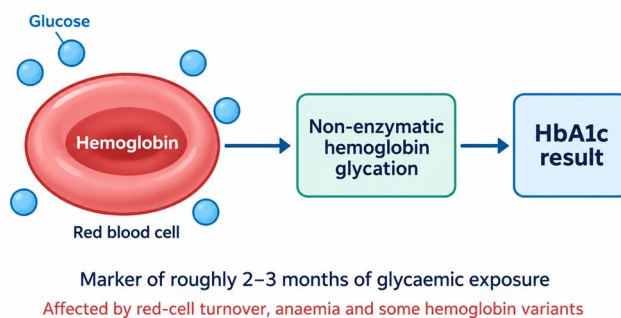


Fig. 1. Original schematic of hemoglobin glycation and interpretation of HbA1c [1].

B. Pattern Recognition for Hospital Risk

Clinical risk arises from interactions among recent utilization, age, comorbidity burden, medications, admission characteristics, discharge disposition, and biochemical evidence of dysglycemia. A multivariable classifier estimates the conditional probability of an outcome from a feature vector and allows threshold-based prioritization. However, a predictor can be statistically associated with readmission without representing an intervention target. In particular, prior utilization and number of laboratory tests can be proxies for severity, access, or documentation intensity.

C. Relevant Literature

TABLE I. COMPARATIVE LITERATURE REVIEW

Study	Method/sample	Finding/limitation
Strack et al. 2014 [3]	HbA1c test ordering; ~70,000 encounters	Testing in 18.4%; relation to readmission varies by diagnosis.
Srinivasan et al. 2020 [4]	Random forest; UCI hospital data	Maximum ROC-AUC 0.684; reported F1 52.07%.
Emi-Johnson & Nkrumah 2025 [5]	LR, RF, XGBoost, DNN; 101,766 encounters	XGBoost ROC-AUC 0.667; class-1 recall remained low.
Salim & Ibrahim 2026 [6]	Nested CV; 4 models plus stacking	Stacking ROC-AUC 0.665; reports calibration and utility.
ADA (2026) [7]	2026 inpatient diabetes clinical guidance	Recommends HbA1c assessment when absent in prior 3 months; emphasizes transition care.
Collins et al. (2024) [8]	TRIPOD+AI prediction-model reporting	27-item reporting framework across statistical and ML models.
Moons et al. (2025) [9]	PROBAST+AI appraisal framework	Evaluates prediction-model quality, risk of bias, and applicability.
Stevens et al. (2020) [10]	Clinical ML reporting recommendations	Highlights transparent datasets, validation, and reproducibility.

Strack et al. [3] analyzed HbA1c measurement and early readmission using nearly 70,000 eligible hospital encounters; HbA1c measurement occurred in 18.4% of encounters, and the adjusted association depended on primary diagnosis. Srinivasan et al. [4] studied random forests for UCI diabetes readmission prediction and reported a maximum ROC-AUC of 0.684 and an F1-score of 52.07%. Emi-Johnson and Nkrumah [5] compared logistic regression, random forest, XGBoost, and a deep neural network on 101,766 encounters; XGBoost had the highest reported ROC-AUC (0.667). Salim and Ibrahim [6] evaluated nested cross-validation, calibration, and clinical utility, with a stacking ROC-AUC of 0.665 under nested validation. These findings motivate a literature-informed experimental protocol; they are not original results of this manuscript.

The studies summarized in Table I use different analyses, model families, and validation choices; therefore, the reported discrimination values should be viewed as context rather than a direct ranking. A recurring research gap is insufficient separation between A1c measurement behavior and the A1c result, together with uncertainty about validation leakage, probability calibration, minority-class sensitivity, and temporal transportability [3]–[6].

The broader evidence base supports stronger standards for clinical deployment and reporting. The 2026 ADA inpatient-care guidance links HbA1c availability to hospital management and discharge transition planning [7]. TRIPOD+AI provides reporting requirements for clinical prediction models [8], while PROBAST+AI offers a structured basis for evaluating model quality, bias, and applicability [9]. Stevens et al. [10] also recommend transparent reporting of data preparation, evaluation, and reproducibility in clinical machine learning.

IV. MATERIALS AND DATA CHARACTERISTICS

A. Data Source and Population

The study framework is based on the Diabetes 130-US Hospitals for Years 1999–2008 dataset, available from the UCI Machine Learning Repository [2]. The collection contains 101,766 inpatient encounters at 130 United States hospitals and integrated delivery networks. Eligible records include hospitalized individuals with diabetes who underwent laboratory examinations, received medication, and stayed 1 to 14 days. The data include administrative, demographic, diagnostic, medication, and utilization attributes. They represent historical inpatient care, not a contemporary prospective cohort.

TABLE II. DATASET CHARACTERISTICS

Attribute	Description
Period	1999–2008
Encounter count	101,766
Hospitals	130 U.S. hospital settings
HbA1c variable	Categorical result; includes “None”
Outcome field	Readmitted: <30, >30, NO
Inputs	Demographics, diagnoses, care use, meds
Cautions	Missingness; repeat patients [2]

B. Target Definition

For the primary task, define $y = 1$ if the readmitted field is “<30” and $y = 0$ if it is “>30” or “NO.” This outcome represents early recorded readmission rather than all future readmissions. A secondary analysis may retain the original three outcome classes. Salim and Ibrahim [6] reported 11,357 encounters (11.16%) coded as readmitted within 30 days in the unfiltered dataset; the count and proportion must be recalculated after eligibility restrictions.

C. Feature Domains

Proposed input domains include: age band, race and gender (subject to fairness assessment), admission type and source, discharge disposition, length of stay, physician specialty, number of laboratory procedures and medications, outpatient/emergency/inpatient encounters, primary and secondary diagnoses, insulin/diabetes-drug indicators, medication changes, and the recorded A1Cresult category. The identifiers encounter_id and patient_nbr must be kept out of the predictive feature set; patient_nbr remains necessary for grouped partitioning.

V. PROPOSED METHODOLOGY

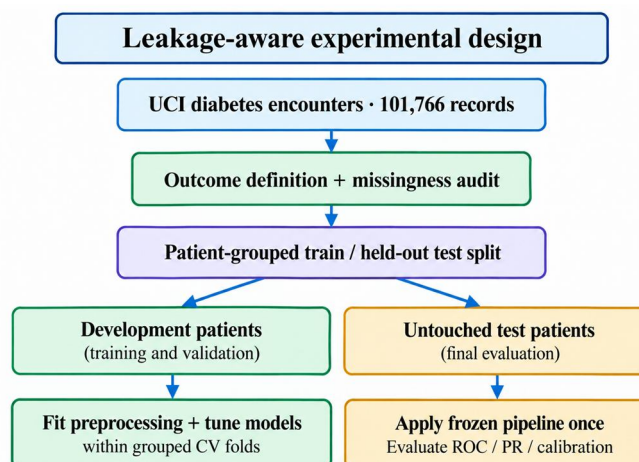


Fig. 2. Proposed patient-separated machine-learning workflow; all fitted transformations are learned on development data only.

A. Study Design and Data Audit

This is a retrospective, methods-oriented secondary-data study proposal. The primary reproducibility unit is the hospital encounter, with patient grouping applied during validation. An audit should report the starting number of encounters, eligibility exclusions, feature-level missingness, duplicate encounter identifiers, repeated-patient frequency, endpoint prevalence, and subgroup counts. Placeholder tokens (notably “?”) should be standardized as missing values without collapsing them into clinically meaningful categories.

B. Preprocessing and Feature Engineering

A prespecified preprocessing pipeline should distinguish numerical, ordered, and nominal features. Numeric laboratory/test counts may be median-imputed and, for linear or margin-based models, standardized. Group rare nominal categories and represent them with one-hot encoding. Missing categorical observations can be assigned an explicit unknown level where informative, but this choice should be evaluated through sensitivity analyses. Highly sparse attributes should be retained only when their interpretation and imputation are credible. Aggregate diagnosis codes into clinically defensible families, and document all mappings.

HbA1c is recorded categorically rather than as a continuous percentage. The original labels should be preserved in analysis: values indicating measurement absent must be distinguished from observed normal and elevated categories. For an ablation analysis, models with and without A1Cresult can quantify incremental predictive information; the same patients, folds, and tuning budget must be used for both runs. The workflow must not construct a numerical HbA1c value from a category midpoint and then imply laboratory precision.

C. Training, Validation, and Leakage Prevention

Split data by patient_nbr into a development partition and an untouched test partition, so no individual appears in both. Within development data, use grouped cross-validation to tune model hyperparameters. Where practical, stratification should accompany grouping to maintain a similar positive-outcome prevalence.

All imputation, feature selection, category grouping based on frequency, scaling, and resampling must be fitted only to the relevant training fold and then applied to its validation fold. Using the full dataset to learn these transformations before splitting would contaminate evaluation and bias reported performance upward.

Suggested fixed test allocation is 20% of distinct patients, subject to preserving a reasonable minority-class count. Random seeds, software versions, folds, eligible records, and decision thresholds should be retained. Class weighting is the first imbalance-handling comparison; resampling methods, if assessed, must occur only inside training folds. Confidence intervals should be estimated using bootstrap resampling at patient level, rather than treating correlated encounters as independent observations.

D. Candidate Models

Five model families support a meaningful complexity ladder. Regularized logistic regression provides a coefficient-based baseline. A pruned decision tree supplies human-readable if-then splits. A random forest offers nonlinear interaction modeling with bagging. A support vector machine can represent nonlinear margins after suitable scaling. Finally, XGBoost or a comparable gradient-boosted tree framework can model structured interactions and missing-value patterns. Hyperparameters should be specified before inspection of the final test data; computationally intensive models should be justified against a simple baseline.

TABLE III. PRESPECIFIED MODEL COMPARISON

Model	Key parameters	Role
Logistic regression	Penalty; C	Interpretable baseline
Decision tree	Depth; min leaf	Simple rules
Random forest	Trees; leaf; max feat.	Nonlinear ensemble
SVM	Kernel; C; gamma	Margin comparator
Gradient boosting	Depth; rate; estimators	Advanced tabular model

E. Performance Measures and Decision Thresholds

Discrimination will be assessed using ROC-AUC and area under the precision-recall curve (PR-AUC), with the latter especially relevant for the minority outcome. At an explicitly chosen operating threshold, report sensitivity, specificity, positive predictive value, negative predictive value, F1 score, false alerts per 100 encounters, and confusion-matrix counts. Probability quality should be assessed with a calibration plot, calibration intercept and slope when estimable, and Brier score. An operational threshold must be chosen on development data according to a defined clinical tradeoff, never by maximizing test-set performance. On imbalanced binary outcomes, precision-recall displays supplement ROC analysis by making positive-class performance more visible [11]. Calibration needs separate assessment because discrimination alone does not establish that predicted probabilities correspond to observed risks [12].

For a probability estimate p_i and observed outcome y_i , the Brier score is $BS = (1/N) \sum_i (p_i - y_i)^2$. Precision is $TP / (TP + FP)$, recall is $TP / (TP + FN)$, and $F1 = 2 \cdot \text{precision} \cdot \text{recall} / (\text{precision} + \text{recall})$. Accuracy may be provided for completeness but is not an acceptable sole measure when about one in nine encounters is a positive event.

F. Explainability, Fairness, and Sensitivity Analysis

Global permutation importance and carefully interpreted local SHAP explanations can identify model associations and possible administrative shortcuts. They do not demonstrate causality. Evaluate ROC-AUC, sensitivity, false-positive rates, and calibration by age band, recorded sex/gender, and race where sample size and ethical constraints allow. Predefine analyses restricted to measured A1c, removal of A1Cresult, first encounter per patient, and omission of variables whose availability at prediction time is uncertain. All subgroup estimates should include sample sizes and uncertainty. SHAP formalizes additive feature-attribution explanations [13], but inherently interpretable alternatives deserve consideration in high-stakes clinical applications [14]. Explanation quality must not be mistaken for demonstrated reliability, validity, or patient benefit [15].

VI. RESULTS: DOCUMENTED EVIDENCE AND REPORTING PLAN

A. Documented Characteristics

No independently executable model training, held-out predictions, or confusion matrices were provided with the source material. Consequently, the present paper does not claim new model performance. Verified dataset-level evidence includes 101,766 historical encounters, categorical A1Cresult, prior utilization variables, and the recorded readmission endpoint [2]. Published work reports 11,357 early-readmission encounters (11.16%) in the full dataset [6].

B. Previously Published Benchmark Evidence

Figure 3 depicts selected published discrimination estimates. Srinivasan et al. [4] reported a maximum random-forest ROC-AUC of 0.684. In the 2025 comparison of Emi-Johnson and Nkrumah [5], logistic regression yielded 0.642, random forest 0.630, and XGBoost 0.667. Salim and Ibrahim [6] reported an ROC-AUC of 0.665 for a stacking ensemble with nested cross-validation. These figures are neither outputs of this paper nor strictly comparable across protocols, cohorts, or timepoints.

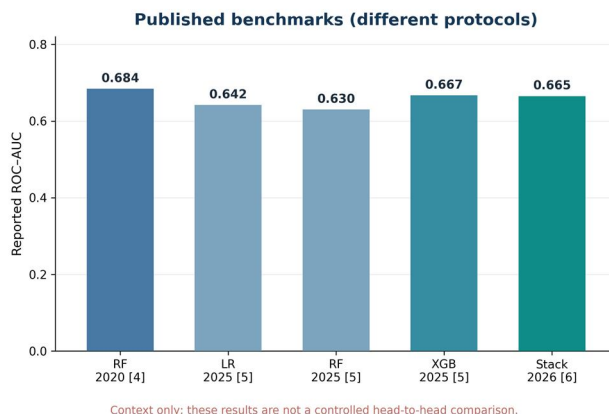


Fig. 3. Published ROC-AUC benchmarks [4]–[6]; different study protocols preclude direct model ranking.

TABLE IV. EVIDENCE STATUS AND REPRODUCIBILITY

Evidence	Source	Status
101,766 encounters	UCI [2]	Dataset characteristic
Readmitted <30: 11.16%	Salim & Ibrahim [6]	Full dataset; recheck after filters
RF AUC 0.684	Srinivasan [4]	Previously published
XGB AUC 0.667	Emi-Johnson [5]	Previously published
New ROC/PR-AUC	This paper	Not yet measured
Calibration & fairness	This paper	Planned analyses

C. Results Tables for Empirical Completion

Once a complete experiment is executed, report the eligible encounter counts and class prevalence for development and test partitions; model hyperparameters and fold assignments; cross-validation means and confidence intervals; untouched-test ROC-AUC and PR-AUC; chosen-threshold operating measures; Brier score and calibration plots; subgroup performance; and A1Cresult ablation effects. An explanation or importance chart should be accompanied by documentation of the prediction timestamp and possible leakage features. No placeholder should be silently converted into a numerical result.

VII. DISCUSSION

The proposed design positions HbA1c as an informative clinical feature rather than as a sufficient descriptor of diabetes-related risk. Historical laboratory results may contribute to risk classification, yet recent hospital use, comorbidity, admission source, medication changes, and discharge context also matter. The key question is not whether an algorithm can recover known diagnostic HbA1c thresholds, but whether combining the recorded A1c category with other routinely recorded variables increases robust prediction of early readmission. Missing HbA1c tests introduce a clinically meaningful complication. Whether clinicians order a test may depend on admission circumstances, resources, and clinical judgment. Treating unmeasured HbA1c as a normal finding would conflate measurement process with underlying glycemia.

Results should therefore distinguish missingness indicators, observed A1c levels, and reduced-cohort analyses. Likewise, repeated encounters can cause inflated model performance if the same patient appears in development and test data. Patient-group-aware splitting is consequently a necessary methodological feature, not merely a preference.

A hospital decision-support model must balance missed high-risk individuals against unnecessary alerts. PR-AUC and positive predictive value relate closely to workload implications, whereas calibration determines whether a displayed risk probability can be interpreted quantitatively. Inference from a historical US inpatient dataset cannot directly establish performance across current treatment pathways or other health systems. Any application beyond research would require prospective evaluation, clinical safety review, subgroup fairness assessment, and local recalibration. Decision-curve analysis is a possible extension for quantifying net clinical benefit over plausible intervention thresholds [16].

From prediction to clinical review

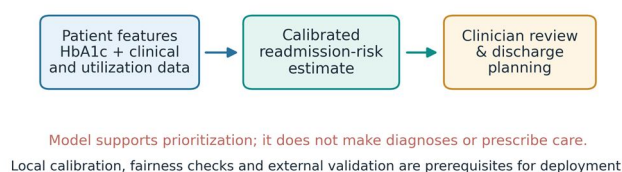


Fig. 4. Proposed clinical decision-support interpretation, conditional on external validation.

VIII. LIMITATIONS AND RESEARCH ETHICS

The underlying encounters were collected between 1999 and 2008. Changes in medicines, practice standards, care coordination, disease prevalence, and recording systems limit generalizability. The HbA1c field is not a continuous laboratory assay result, and important drivers of readmission—social support, affordability, health literacy, renal function detail, continuous glucose monitoring, and post-discharge care—may be incomplete or unavailable. Dataset missingness may be nonrandom. Readmission can also occur outside the original hospital network and may therefore be incompletely ascertained.

The study proposes a secondary analysis using a public, de-identified repository and describes no new participant recruitment or clinical intervention. Before real deployment or linkage with patient-level institutional data, local governance, privacy review, dataset terms, and applicable ethics requirements must be assessed. Predictive performance alone does not justify automated treatment choices. Any implementation pathway must include fairness and clinical oversight. Analyses of healthcare risk algorithms have shown that a seemingly neutral prediction target can produce inequitable outcomes; subgroup assessment should therefore consider the meaning of the target as well as model error rates [17].

IX. FUTURE WORK

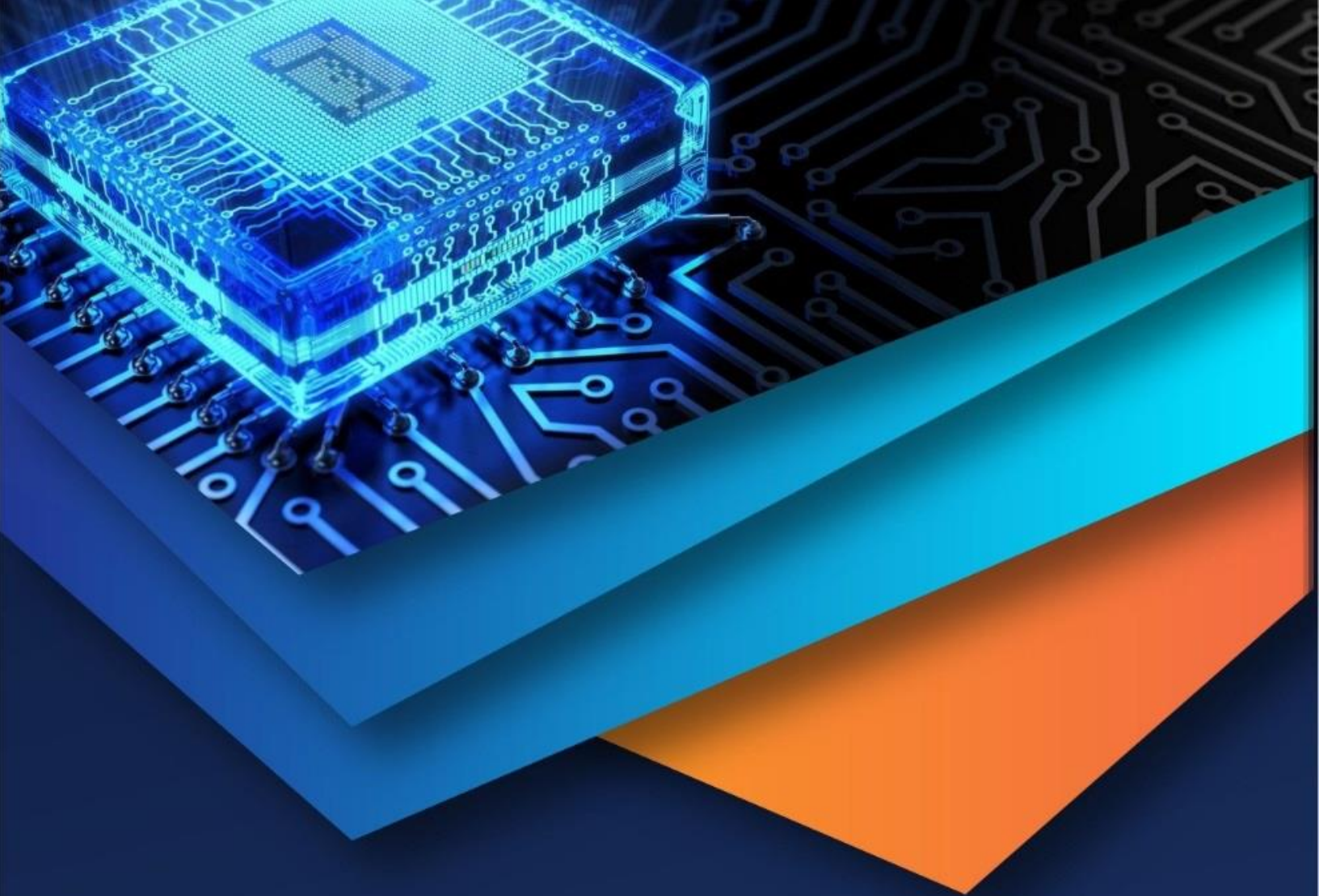
The next step is to obtain and analyse the raw UCI files using patient-separated partitions and an auditable, leakage-safe pipeline. Researchers should release the feature dictionary, eligibility decisions, random seeds, software versions, tuning search space, and confidence intervals. A1Cresult ablation can examine incremental benefit beyond hospital-utilization features; follow-up work should test clinical alert burden and threshold sensitivity. Contemporary external datasets with quantitative HbA1c and additional laboratory and social factors are necessary to evaluate changes in performance over time. If a validated model eventually progresses to an interventional evaluation, its trial reporting should follow relevant AI-specific guidance, including CONSORT-AI [18].

X. CONCLUSION

HbA1c-informed pattern recognition offers a plausible framework for organizing diabetes hospital records into clinically interpretable risk profiles. The UCI Diabetes 130-US Hospitals dataset provides a substantial historical foundation for comparing transparent baselines and nonlinear machine-learning methods, but it does not by itself establish a deployable predictor. A defensible analysis must separate A1c measurement from observed category, prevent repeated-patient leakage, quantify minority-class detection, assess probability calibration, and validate across patient groups. The present paper contributes a reproducible study design and evidence synthesis while explicitly reserving original performance claims for empirical execution. Its anticipated role is decision support for clinical review and discharge planning, never an autonomous substitute for professional judgment.

REFERENCES

- [1] American Diabetes Association Professional Practice Committee, "2. Diagnosis and Classification of Diabetes: Standards of Care in Diabetes—2026," *Diabetes Care*, vol. 49, suppl. 1, pp. S27–S49, 2026, doi: 10.2337/dc26-S002.
- [2] J. Clore, K. Cios, J. DeShazo, and B. Strack, "Diabetes 130-US Hospitals for Years 1999–2008," UCI Machine Learning Repository, 2014, doi: 10.24432/C5230J.
- [3] B. Strack et al., "Impact of HbA1c measurement on hospital readmission rates: Analysis of 70,000 clinical database patient records," *BioMed Research International*, vol. 2014, Art. no. 781670, 2014, doi: 10.1155/2014/781670.
- [4] S. M. Srinivasan, Y. F. Paat, P. Halls, R. Kalule, and T. E. Harvey, "Predicting re-admission to hospital for diabetes treatment: A machine learning solution," *International Journal of Computational Biology and Drug Design*, vol. 13, nos. 5–6, pp. 539–554, 2020, doi: 10.1504/IJCDD.2020.113823.
- [5] O. G. Emi-Johnson and K. J. Nkrumah, "Predicting 30-Day Hospital Readmission in Patients With Diabetes Using Machine Learning on Electronic Health Record Data," *Cureus*, vol. 17, no. 4, Art. no. e82437, Apr. 2025, doi: 10.7759/cureus.82437.
- [6] S. S. Salim and A. A. Ibrahim, "A Machine Learning Approach for Predicting 30-Day Hospital Readmission in Patients with Diabetes," *Healthcare*, vol. 14, no. 9, Art. no. 1185, Apr. 2026, doi: 10.3390/healthcare14091185.
- [7] American Diabetes Association Professional Practice Committee for Diabetes, "16. Diabetes Care in the Hospital: Standards of Care in Diabetes—2026," *Diabetes Care*, vol. 49, suppl. 1, pp. S339–S355, 2026, doi: 10.2337/dc26-S016.
- [8] G. S. Collins et al., "TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods," *BMJ*, vol. 385, Art. no. e078378, 2024, doi: 10.1136/bmj-2023-078378.
- [9] K. G. M. Moons et al., "PROBAST+AI: An updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods," *BMJ*, vol. 388, Art. no. e082505, 2025, doi: 10.1136/bmj-2024-082505.
- [10] L. M. Stevens, B. J. Mortazavi, R. C. Deo, L. Curtis, and D. P. Kao, "Recommendations for reporting machine learning analyses in clinical research," *Circ. Cardiovasc. Qual. Outcomes*, vol. 13, no. 10, Art. no. e006556, 2020, doi: 10.1161/CIRCOUTCOMES.120.006556.
- [11] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," *PLoS ONE*, vol. 10, no. 3, Art. no. e0118432, 2015, doi: 10.1371/journal.pone.0118432.
- [12] B. Van Calster et al., "Calibration: The Achilles heel of predictive analytics," *BMC Med.*, vol. 17, Art. no. 230, 2019, doi: 10.1186/s12916-019-1466-7.
- [13] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [14] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nat. Mach. Intell.*, vol. 1, pp. 206–215, 2019, doi: 10.1038/s42256-019-0048-x.
- [15] M. Ghassemi, L. Oakden-Rayner, and A. L. Beam, "The false hope of current approaches to explainable artificial intelligence in health care," *Lancet Digit. Health*, vol. 3, no. 11, pp. e745–e750, 2021, doi: 10.1016/S2589-7500(21)00208-9.
- [16] A. J. Vickers and E. B. Elkin, "Decision curve analysis: A novel method for evaluating prediction models," *Med. Decis. Making*, vol. 26, no. 6, pp. 565–574, 2006, doi: 10.1177/0272989X06295361.
- [17] Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations," *Science*, vol. 366, no. 6464, pp. 447–453, 2019, doi: 10.1126/science.aax2342.
- [18] X. Liu et al., "Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension," *Nat. Med.*, vol. 26, pp. 1364–1374, 2020, doi: 10.1038/s41591-020-1034-x.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)