

Article

Not peer-reviewed version

Predicting Iron Deficiencies Using Routine Complete Blood Cell Count Parameters: A Machine Learning Approach and Evaluation

[Davide Negrini](#) , [Laura Pighi](#) , [Simone Mignolli](#) , [Gian Luca Salvagno](#) , [Giuseppe Lippi](#) *

Posted Date: 2 April 2026

doi: 10.20944/preprints202604.0073.v1

Keywords: machine learning; prediction algorithms; blood cell count; iron deficiencies



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Predicting Iron Deficiencies Using Routine Complete Blood Cell Count Parameters: A Machine Learning Approach and Evaluation

Davide Negrini, Laura Pighi, Simone Mignolli, Gian Luca Salvagno and Giuseppe Lippi *

Section of Clinical Biochemistry, University of Verona, Verona, Italy

* Correspondence: giuseppe.lippi@univr.it

Abstract

Background/Objectives: Iron deficiency remains a prevalent condition, needing specific laboratory tests for diagnosis. This study aimed to evaluate whether routine complete blood cell count (CBC) parameters can be used within a machine learning framework to predict iron deficiency, potentially optimizing laboratory test utilization. **Methods:** A retrospective dataset of outpatients (2023–2026) undergoing both CBC and iron testing was analyzed. Iron deficiency was defined using sex-specific thresholds for ferritin and transferrin saturation. After cleaning data and excluding incomplete records, demographic variables and CBC indices were tested as potential predictors. The dataset was split into training and test sets with stratified sampling. Multiple supervised machine learning models, including logistic regression, decision tree, random forest, XGBoost, support vector machine, k-nearest neighbors, and Naive Bayes, were trained. Hyperparameter tuning and model selection were performed using repeated stratified 10-fold cross-validation, optimizing the area under the curve (AUC). Model performance was assessed by AUC, sensitivity, and specificity, and validated on an independent test set. **Results:** All models demonstrated predictive capability using CBC parameters alone. Ensemble methods, especially random forest and XGBoost, reached the best performance (AUC values of 0.80–0.87 for ferritin and 0.85–0.96 for transferrin saturation). Sensitivity and specificity were balanced, supporting clinical screening applicability. Results were maintained across validation and confirmed in the test set. Prediction of transferrin saturation showed slightly higher accuracy than ferritin. Feature importance analysis identified MCV, MCH, and RDW as key predictors. **Conclusions:** CBC-based machine learning models can reliably identify subjects with iron deficiency, supporting subsequent, more targeted analyses.

Keywords: machine learning; prediction algorithms; blood cell count; iron deficiencies

1. Introduction

Anemia is a significant public health problem worldwide, affecting both high- and low-income countries. In Europe, iron deficiency remains one of the most common nutritional deficiencies, affecting a substantial proportion of the population [1]. Older adults are at increased risk of iron deficiency, as aging is frequently associated with inadequate dietary intake, reduced gastrointestinal absorption often related to proton pump inhibitor use, and chronic blood loss due to angiodysplasia, gastrointestinal malignancies, or concomitant antithrombotic therapy [2].

Laboratory investigations for the diagnosis of iron deficiency and anemia include a variety of tests, such as a complete blood count (CBC), serum iron, ferritin, transferrin, and transferrin saturation. The diagnostic workup of anemia should follow a stepwise approach in which the CBC serves as the initial investigation and guides further laboratory testing, whereas iron studies, including ferritin and transferrin parameters, are performed only when anemia or microcytosis is detected, to ensure appropriate and cost-effective use of laboratory resources [3,4].

CBC parameters are inexpensive, widely available, and routinely used in clinical practice. Importantly, erythrocyte indices such as mean corpuscular volume (MCV), mean corpuscular hemoglobin, and red blood cell distribution width (RDW) reflect alterations in erythropoiesis and may hence provide indirect evidence of iron-restricted erythropoiesis even before biochemical abnormalities become evident.

Machine learning methods have been applied in laboratory medicine to analyze laboratory data and to support diagnostic decision-making [5]. The large volume of laboratory data generated in clinical practice has allowed the development of models for earlier detection and more efficient use of laboratory resources. Previous studies have shown that machine learning can be integrated into laboratory medicine to support diagnostics and improve test utilization. Laboratory data are highly standardized and quantitative, which makes them suitable for machine learning applications. These approaches can improve test interpretation, support risk stratification, and contribute to more appropriate use of laboratory tests [6]. The value of this approach has been demonstrated in different diagnostic settings, where machine learning models based on routinely available laboratory and biochemical parameters have shown good performance in disease detection and classification [7,8]. The same analytical approach has also been applied to routine laboratory data to identify conditions that may not be easily detected by conventional procedures, including laboratory errors and preanalytical interferences such as wrong-blood-in-tube events and contamination from intravenous fluids [9,10].

These studies clearly show that laboratory data obtained in clinical practice can reveal patterns that are hardly detectable with conventional analytical approaches. Based on these considerations, the aim of our study was to evaluate the feasibility of a machine learning approach using only complete blood count parameters to identify subjects at risk of iron deficiency.

2. Materials and Methods

2.1. Dataset and Targeted Markers

We extracted from the Laboratory Information System (LIS) all outpatients referred to our Laboratory who had, in the same blood collection, CBC, iron, ferritin, and transferrin requests, between 1st January 2023 and 31st January 2026. We excluded all patients with incomplete results or information (so no imputation was needed) and kept only patients aged 10-99 (inclusive). CBC tests have been performed using a Sysmex XN hematological system (Sysmex Corp., Kobe, Japan), while iron, ferritin, and transferrin in plasma were measured on Cobas c701 with proprietary kits (Roche Diagnostics AG, Risch-Rotkreuz, Switzerland).

The final dataset was completely anonymized and included only the following columns: age (years), gender (Male or Female), red blood cell (RBC, $\times 10^{12}/L$) count, hemoglobin (HB, g/L), hematocrit (HT, L/L), MCV (fL), mean corpuscular hemoglobin (pg), mean corpuscular hemoglobin concentration (MCHC, g/L), RDW (%), plasma iron concentration ($\mu\text{mol}/L$), plasma ferritin concentration ($\mu\text{g}/L$), plasma transferrin concentration (g/L). Transferrin saturation in % has been calculated using the following formula $[(\text{iron} \times 5.5845) / (\text{transferrin} \times 138.9)]$ [11,12].

Targeted biomarkers for iron deficiency identification were (A) low ferritin concentration, considered as lower than 30 $\mu\text{g}/L$ in females and 50 $\mu\text{g}/L$ in males, and (B) low transferrin saturation, considered as lower than 15% in females and 20% in males.

2.2. Pipeline Implementation

A supervised machine learning approach was adopted to develop and compare binary classification models for predicting two distinct iron-deficiency biomarkers using routine CBC parameters alone. Two separate but methodologically identical analysis pipelines were implemented, one for each target biomarker.

For each pipeline, only the parameters age, gender, RBC, HB, HT, MCV, MCH, MCHC, and RDW were given as input, removing the values of the iron metabolism biomarkers, and including only a binary label as the prediction target (presence or absence of the low value).

To implement the desired pipeline, we created two R scripts (R Core Team (2025). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria) differing only in the biomarker targeted, written in Visual Studio Code (Microsoft Corp., Redmond, WA, USA) and reviewed for optimization using GitHub Copilot with Claude Opus 4.6 (Anthropic PBC, San Francisco, CA, USA).

The R script was then run on R 4.5.2 (2025-10-31) with RStudio Pro 2025.09.1 (Posit Software PBC, Boston, MA; USA) on a x64-pc-linux-gnu platform.

First of all, the dataset was split into a training set (70%) and a held-out test set (30%) using stratified random sampling via the `createDataPartition` function of “`caret`” package [13], which preserves the original class distribution in both subsets. A constant random seed was used for all analyses to ensure comparable splits.

Most models (all except XGBoost and Naive Bayes) were trained and tuned using a unified framework via the “`caret`” package. A repeated stratified 10-fold cross-validation scheme (3 repeats, 30 total resampling iterations) was applied on the training set. The optimization metric was the area under the receiver operating characteristics (ROC) curve (AUC). Cross-validated estimates of AUC, sensitivity, and specificity (mean +/- standard deviation across folds) were recorded for all models using the `twoClassSummary` function. Class probability estimation was enabled for all classifiers to support ROC analysis.

2.3. Classifiers Evaluation

Nine classifiers spanning different algorithmic families were trained and evaluated:

(a) Logistic Regression (`caret` method “`glm`”, predictions obtained by applying a 0.5 probability threshold to the predicted probabilities),

(b) Basic Decision Tree (`caret` method “`rpart`”, fixed complexity parameter `cp` = 0.01),

(c) Optimized Decision Tree (`caret` method “`rpart`”, with hyperparameter tuning of the complexity parameter (`cp`) over a grid ranging from 0.001 to 0.05 in steps of 0.005. The optimal `cp` was selected as the value maximizing the cross-validated AUC),

(d) Basic Random Forest (`caret` method “`rf`”, with 500 trees and fixed `mtry` = 5),

(e) Tuned Random Forest: (`caret` method “`rf`”, with 500 trees with hyperparameter tuning of `mtry` (values: 2, 3, 4, 5, 6) to maximize cross-validated AUC),

(f) XGBoost (Extreme Gradient Boosting) (trained using “`xgboost`” package, with a custom repeated cross-validation procedure (10-fold x 3 repeats) implemented using `createMultiFolds`, ensuring methodological consistency with the other models; grid search was performed over the following hyperparameters: number of boosting rounds (`nrounds`: 50, 100, 150), maximum tree depth (`max_depth`: 3, 6), learning rate (`eta`: 0.1, 0.3)),

(g) Support Vector Machine (SVM) with a radial basis function (`caret` method “`svmRadial`”, trained with automatic tuning of the `sigma` and `cost` (C) parameters with `tuneLength` = 5; features were centered and scaled as part of the `caret` preprocessing pipeline),

(h) k-Nearest Neighbors (k-NN): A k-NN classifier (`caret` method “`knn`”, trained with hyperparameter tuning of `k` over odd values from 3 to 25, optimal `k` selected by maximizing cross-validated AUC on the training set; features were centered and scaled as part of the `caret` preprocessing pipeline),

(i) Naive Bayes, assuming Gaussian feature distributions (trained using the “`e1071`” package, with a custom repeated cross-validation procedure (10-fold x 3 repeats), was implemented using `createMultiFolds`, ensuring methodological consistency with the other models).

Each model was evaluated on the held-out test set (30% of data, never used during training or tuning). The following metrics were computed from the confusion matrix with the positive class set to “1”: Accuracy, Balanced Accuracy (average of sensitivity and specificity), Sensitivity, Specificity,

Positive Predictive Value, Negative Predictive Value, F1-score (harmonic mean of precision and sensitivity).

For all models providing class probability estimates, ROC curves were constructed on the test set and the AUC was computed using the “pROC” package [14]. The 95% confidence interval for each AUC was estimated using the DeLong method [15].

To determine whether the discriminative performance of the best-performing model was statistically significantly different from each of the other models, pairwise DeLong tests (roc.test function, “pROC” package) were performed on the test-set ROC curves. Statistical significance was assessed at alpha = 0.05. The best model was identified as the one with the highest test-set AUC and served as the reference for all pairwise comparisons.

3. Results

The final dataset had a total of 32437 rows with complete data about the 9 prediction factors (age, gender, RBC, HB, HT, MCV, MCH, MCHC, RDW) and the selected biomarkers, of which 15917 males and 16520 females; age range 10-99 years, median age 65 (25°-75° percentile: 48-78). With a 70:30 split, we had 22,707 rows in the training set and 9,730 in the test set.

For the low ferritin, we had a target distribution of 18093 negative (non-low value, 55.8%), 14344 positive (low value, 44.2%); results of the different classifiers on the holdout test set are reported in Table 1. In Figure 1 a graphical AUC comparison is reported.

Given the low transferrin saturation, we set a target distribution of 24646 negative (non-low value, 76.0%) and 7791 positive (low value, 24.0%). Results of the different classifiers on the holdout test set are reported in Table 2. In Figure 2 a graphical AUC comparison is reported.

Table 1. Results of the different classifiers on the holdout test for the low ferritin target.

Model	Accuracy	Balanced accuracy	Sensitivity	Specificity	PPV	NPV	F1-score	AUC	AUC 95% CI
Logistic Regression	0.6383	0.6196	0.4578	0.7815	0.6242	0.6451	0.5282	0.6634	0.6525 - 0.6742
Basic Decision Tree	0.7681	0.7658	0.7458	0.7859	0.7342	0.7959	0.7399	0.8107	0.8021 - 0.8194
Optimized Decision Tree	0.7947	0.7880	0.7300	0.8460	0.7898	0.7980	0.7587	0.8354	0.8271 - 0.8436
Basic Random Forest	0.7902	0.7851	0.7402	0.8299	0.7753	0.8011	0.7573	0.8500	0.8422 - 0.8578
Tuned Random Forest	0.7941	0.7893	0.7476	0.8310	0.7782	0.8059	0.7626	0.8529	0.8451 - 0.8606
XGBoost	0.8016	0.7962	0.7495	0.8430	0.7910	0.8093	0.7697	0.8584	0.8507 - 0.8660
SVM	0.6689	0.6514	0.5006	0.8023	0.6675	0.6695	0.5721	0.7119	0.7016 - 0.7221
k-NN (tuned as k=25)	0.6563	0.6428	0.5254	0.7601	0.6346	0.6689	0.5749	0.6956	0.6852 - 0.7061
Naive Bayes	0.6405	0.6241	0.4822	0.7660	0.6203	0.6511	0.5426	0.6649	0.6541 - 0.6758

AUC, Area under the curve; PPV, positive predictive value; NPV, negative predictive value.

Table 2. Results of the different classifiers on the holdout test for the low transferrin saturation target.

Model	Accuracy	Balanced accuracy	Sensitivity	Specificity	PPV	NPV	F1-score	AUC	AUC 95% CI
-------	----------	-------------------	-------------	-------------	-----	-----	----------	-----	------------

Logistic Regression	0.7814	0.5898	0.2212	0.9585	0.6274	0.7956	0.3271	0.7496	0.7383 - 0.7610
Basic Decision Tree	0.7764	0.5893	0.2294	0.9493	0.5884	0.7958	0.3300	0.6759	0.6644 - 0.6874
Optimized Decision Tree	0.7978	0.6519	0.3710	0.9328	0.6356	0.8243	0.4685	0.7419	0.7299 - 0.7539
Basic Random Forest	0.8016	0.6591	0.3847	0.9335	0.6463	0.8276	0.4823	0.7773	0.7663 - 0.7882
Tuned Random Forest	0.7996	0.6469	0.3530	0.9408	0.6532	0.8214	0.4583	0.7801	0.7692 - 0.7909
XGBoost	0.8063	0.6650	0.3932	0.9368	0.6631	0.8301	0.4937	0.8013	0.7911 - 0.8115
SVM	0.7947	0.6090	0.2516	0.9663	0.7025	0.8033	0.3705	0.7278	0.7150 - 0.7406
k-NN (tuned as k=25)	0.7937	0.6307	0.3171	0.9444	0.6432	0.8139	0.4248	0.7684	0.7574 - 0.7795
Naive Bayes	0.7503	0.6198	0.3688	0.8708	0.4744	0.8136	0.4150	0.7266	0.7151 - 0.7382

AUC, Area under the curve; PPV, positive predictive value; NPV, negative predictive value.

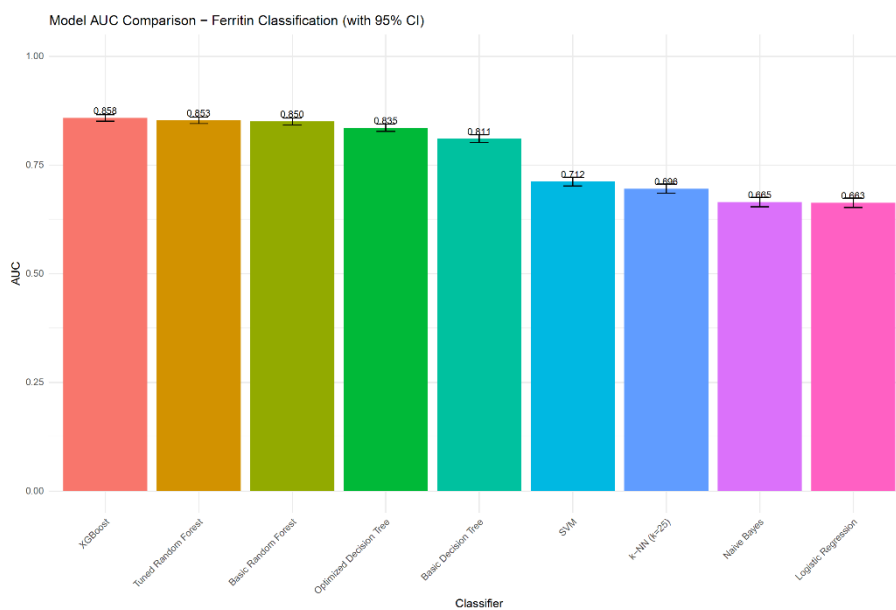


Figure 1. Graphical comparison of AUCs of the different classifiers for the low ferritin target. *AUC, area under the curve.*



Figure 2. Graphical comparison of AUCs of the different classifiers for the low transferrin saturation target. *AUC*, area under the curve.

For both targets, XGBoost achieved the highest AUC, and the model, compared with all others evaluated using the DeLong test, always had a p-value < 0.05; ROC curves are displayed in Figures 3 and 4.

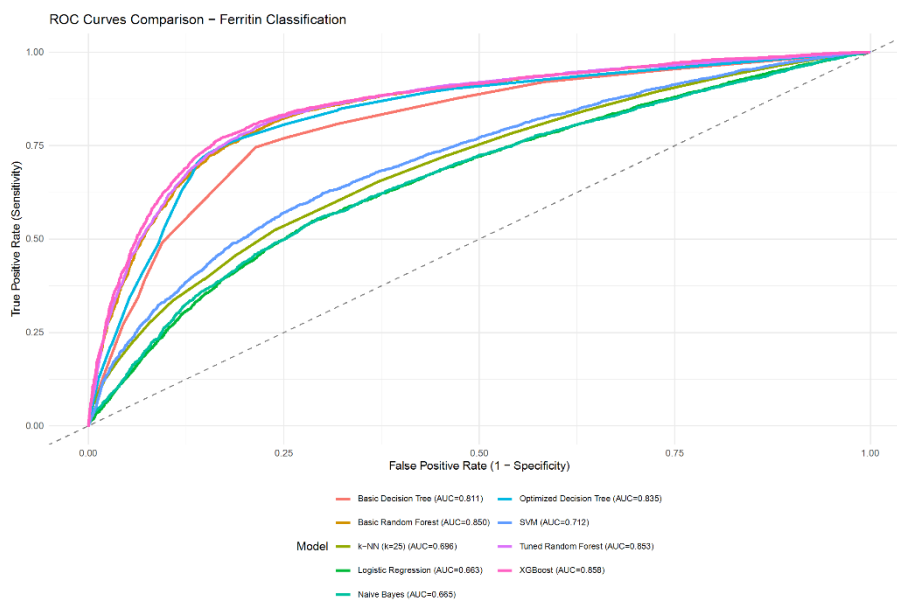


Figure 3. ROC curves of the different classifiers for the low ferritin target. *ROC*, Receiver Operating Characteristics.

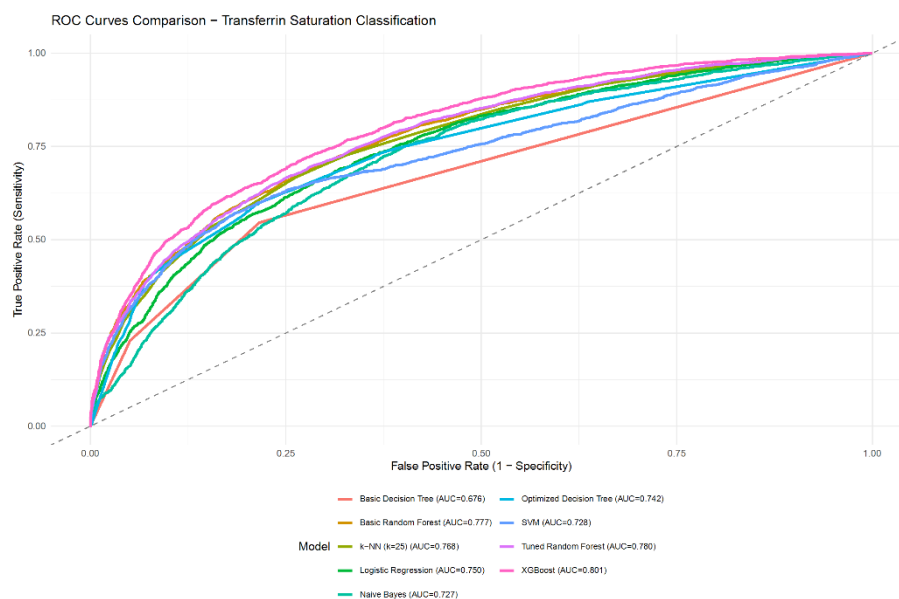


Figure 4. ROC curves of the different classifiers for the low transferrin saturation target. ROC, Receiver Operating Characteristics.

As an additional evaluation, feature importance was extracted from the tuned Random Forest model using the Mean Decrease Gini index, which quantifies the total decrease in node impurity (Gini impurity) attributable to each feature across all trees in the ensemble. Features were ranked in descending order of importance and are presented graphically for the two targets in Figures 5 and 6. For the ferritin target, the most important predictor was MCH; for the transferrin saturation target, it was MCV.

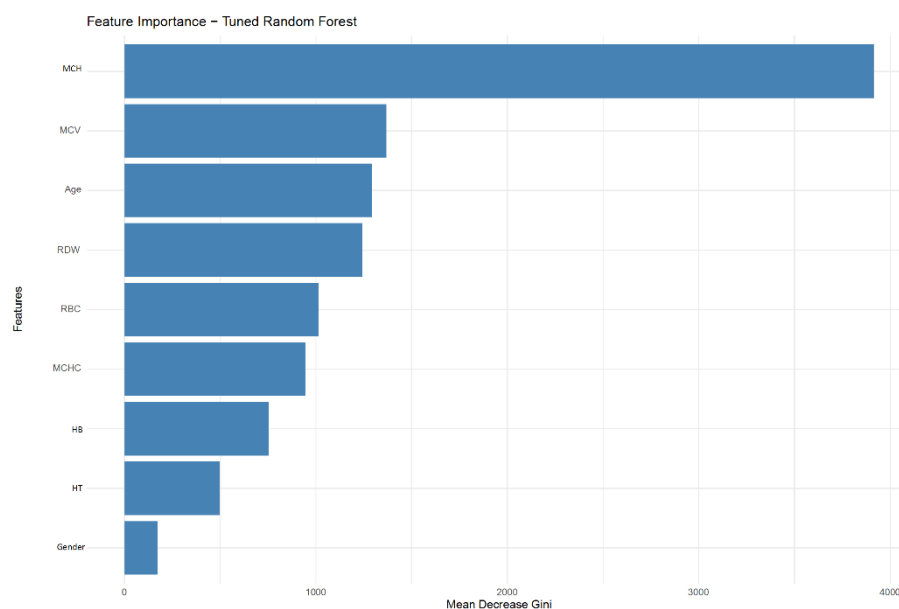


Figure 5. Feature importance extracted from the tuned Random Forest model using the Mean Decrease Gini index for the low ferritin target.

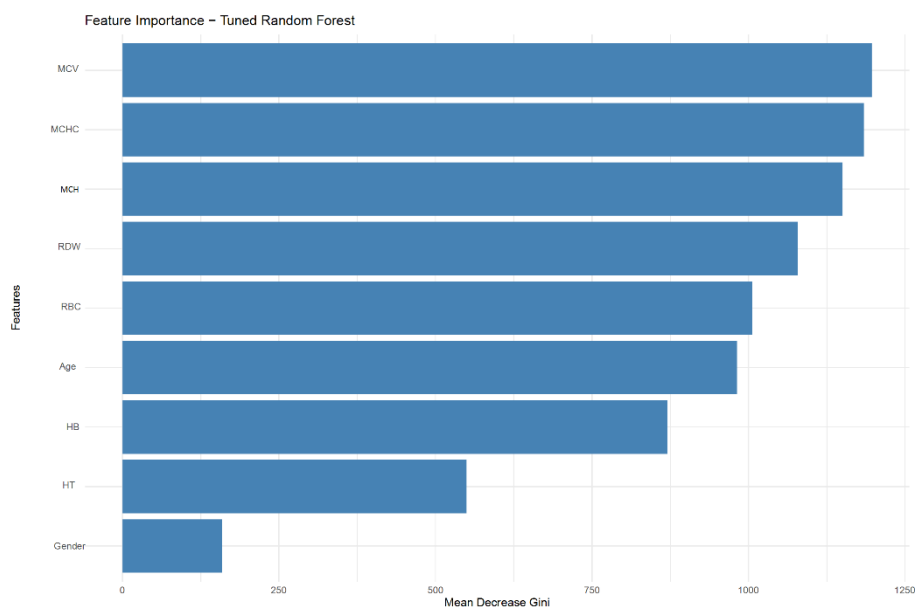


Figure 6. Feature importance extracted from the tuned Random Forest model using the Mean Decrease Gini index for the low transferrin saturation target.

4. Discussion

In recent years, machine learning methods have been applied in laboratory medicine to analyze complex quantitative datasets. Supervised models are trained on predefined laboratory variables to generate classifications or quantitative estimates, using structured input features derived from standardized analytical procedures [16]. The large volume of numerical data generated in laboratory testing provides a suitable basis for developing such models. Within hematological testing, complete blood count parameters constitute a well-standardized, numerically structured set of measurements, making them appropriate for multivariable modeling. Tepakhan and colleagues applied combined decision-tree models to red blood cell indices to differentiate iron deficiency anemia from thalassemia, providing an example of computational models developed exclusively from laboratory variables [17].

In the present study, we evaluated whether complete blood count parameters could predict biochemical evidence of iron deficiency in a large outpatient population. Two parallel supervised classification pipelines were developed to identify low ferritin concentration and low transferrin saturation using exclusively hematological variables as input features. Among the evaluated classifiers, XGBoost achieved the highest AUC for both targets and remained significantly superior in direct ROC comparisons using the DeLong test. For the ferritin-based model, the positive predictive value was 0.79, indicating that a substantial proportion of subjects classified as positive had reduced ferritin concentrations. Conversely, the transferrin saturation model yielded a negative predictive value of 0.83, supporting its ability to reliably exclude reduced transferrin saturation. Considered together, these performance characteristics suggest that the two models may provide complementary information when interpreted within the same laboratory framework, one contributing to rule-in performance and the other to rule-out capability.

The consistently higher discriminative performance of tree-based boosting methods, particularly for the ferritin target, indicates that the association between erythrocyte indices and iron stores is likely characterized by nonlinear interactions that are not fully captured by linear approaches. This interpretation is further supported by the feature importance analysis. For the ferritin model, MCH emerged as the most influential predictor. This finding is biologically coherent and supported by previous evidence demonstrating a significant association between serum ferritin and MCH levels in patients with iron-deficiency anemia [18]. Reduced iron availability directly limits hemoglobin synthesis, leading to a decrease in hemoglobin content per erythrocyte before more

pronounced changes in cell size occur. In this context, MCH may reflect iron-restricted erythropoiesis at an earlier or more quantitatively sensitive stage than other red cell indices. For the transferrin saturation model, MCV ranked highest in importance. This observation is consistent with the progressive reduction in erythrocyte size observed in iron-deficient states, in which sustained impairment of hemoglobin synthesis ultimately leads to microcytosis. The differential ranking of MCH and MCV across the two targets suggests that distinct hematological expressions of iron deficiency may be preferentially captured depending on whether the biochemical reference marker reflects iron stores or circulating iron availability.

Our findings are also in line with those reported by Garduno-Rapp and colleagues, who developed deep learning models for early identification of iron-deficiency anemia using 52 longitudinal laboratory parameters and achieved an AUROC of 0.89 with a gated recurrent unit architecture [19]. Although their approach incorporated a substantially broader panel of laboratory variables and a time-series structure, the overall discriminative performance observed in our study was comparable, despite using only age, sex, and routine complete blood count indices. This comparison underscores that a limited yet biologically coherent hematological dataset may retain considerable predictive information about iron deficiency, even in the absence of extended biochemical panels or longitudinal modeling. Beyond the methodological comparison, this finding also has practical implications. Because the model is based on only age, sex, and routine complete blood count indices, it may be applicable across a broad range of laboratory settings, including those in which ferritin, transferrin, or transferrin saturation are not consistently available. In such circumstances, a single routine analysis may already support the interpretation of the hematological profile and help identify patients with a higher likelihood of iron deficiency, thereby guiding further biochemical assessment when feasible. The limited number of required variables may also facilitate implementation in smaller laboratories and compact hematology platforms.

Recent studies have applied gradient-boosted decision tree models to routinely collected clinical and laboratory variables to identify iron deficiency in subjects without overt anemia [20]. These observations are concordant with the present findings and indicate that biochemical iron depletion may already be reflected in quantitative alterations of hematological indices before hemoglobin levels decline below the diagnostic threshold for anemia. The ability to capture these early changes through multivariable analysis supports the biological plausibility of the approach adopted in this study.

In the present analysis, hyperparameter tuning was explicitly performed to maximize AUC as the primary optimization criterion. Future studies may reasonably explore alternative tuning strategies tailored to specific laboratory applications, for example, prioritizing sensitivity in screening-oriented settings or specificity in confirmatory contexts. The results of the present analysis must be interpreted in light of the specific methodological framework in which the models were developed. Although performance was assessed on a held-out test subset, all data originated from a single institutional setting. Independent validation on external cohorts, ideally generated on different analytical platforms and reflecting heterogeneous patient populations, will be required to determine the stability and broader applicability of the proposed models. It is also essential to clarify that the analytical strategy described herein is not conceived as a replacement for established diagnostic pathways. Iron deficiency remains a condition requiring biochemical confirmation and clinical contextualization. The potential contribution of a model based exclusively on complete blood count parameters lies in its ability to highlight hematological patterns that may warrant targeted second-line investigations, thereby supporting a more focused use of iron studies in routine laboratory practice.

A further aspect deserving consideration is the absence of explicit assessment of inflammatory status. Ferritin is a positive acute-phase reactant and may increase independently of iron stores in the presence of systemic inflammation, while transferrin concentrations typically decline under similar conditions. The lack of inflammatory markers in the dataset may therefore have influenced the classification of certain cases in which alterations in iron-related parameters were driven, at least in part, by inflammatory mechanisms rather than by absolute iron deficiency. Incorporation of indices

such as C-reactive protein or comparable markers could allow a more refined interpretation of iron metabolism parameters in future model iterations.

5. Conclusions

The present study provides a basis for the development of laboratory-based diagnostic support strategies for the identification of iron deficiency. Complete blood count analysis is a widely available, standardized, and low-cost test routinely performed across different clinical settings. The observation that hematological parameters alone provide meaningful discriminatory performance suggests that information derived from routine hematological testing may assist in recognizing patterns consistent with iron depletion, including contexts in which iron studies are not included in the initial test panel, are selectively requested, or are unavailable at the time of evaluation. This may also be relevant in low-resource settings, where a single routine analysis may provide an initial interpretation of the hematological profile and help identify patients who may require further evaluation of iron status. In this context, the complete blood count may help guide the use of dedicated iron studies and enhance the interpretive value of routine laboratory testing. Future integration of this approach into laboratory information systems may facilitate its use in routine clinical practice.

Author Contributions: Conceptualization D.N., L.P., G.L.S., G.L.; software, D.N., S.M.; validation, D.N., S.M.; formal analysis, D.N., L.P.; data curation, D.N., S.M., L.P.; writing—original draft preparation, D.N., L.P.; writing—review and editing, L.P., G.L.S., G.L.; supervision, G.L.S., G.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were not required for this study because it used only aggregated, completely anonymized historical data.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: During the preparation of this manuscript/study, the author(s) used GitHub Copilot with Claude Opus 4.6 for the purposes of code review and optimization. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AUC	Area Under the Curve
CBC	Complete Blood Cell Count
DOAJ	Directory of Open Access Journals
HB	Hemoglobin
HT	Hematocrit
LD	Linear dichroism
LIS	Laboratory Information System
MCH	Mean Corpuscular Hemoglobin
MCHC	Mean Corpuscular Hemoglobin Concentration
MCV	Mean Corpuscular Volume
MDPI	Multidisciplinary Digital Publishing Institute
NPV	Negative Predictive Value
PPV	Positive Predictive Value

RBC	Red Blood Cell
RDW	red blood cell distribution width
ROC	Receiver Operating Characteristics
TLA	Three letter acronym

References

1. Hercberg S, Preziosi P, Galan P. Iron deficiency in Europe. *Public Health Nutrition*. 2001;4(2b):537-545. doi:10.1079/PHN2001139.
2. Girelli D, Marchi G, Camaschella C. Anemia in the Elderly. *Hemasphere*. 2018;2(3):e40. Published 2018 Apr 17. doi:10.1097/HS9.0000000000000040
3. Wang D, Sra M, Glaeser-Khan S, et al. Cost-Effectiveness of Ferritin Screening Thresholds for Iron Deficiency in Reproductive-Age Women. *Am J Hematol*. 2025;100(7):1132-1140. doi:10.1002/ajh.27686.
4. Sholzberg M, Hillis C, Crowther M, Selby R. Diagnosis and management of iron deficiency in females. *CMAJ*. 2025;197(24):E680-E687. Published 2025 Jul 1. doi:10.1503/cmaj.240570.
5. Lippi G, Plebani M. Lights and shadows of artificial intelligence in laboratory medicine. *Adv Lab Med*. 2025;6(1):1-3. Published 2025 Feb 24. doi:10.1515/almed-2025-0024
6. Rabbani N, Kim GYE, Suarez CJ, Chen JH. Applications of machine learning in routine laboratory medicine: Current state and future directions. *Clin Biochem*. 2022;103:1-7. doi:10.1016/j.clinbiochem.2022.02.011
7. Ning W, Wang Z, Gu Y, et al. Machine learning models based on routine blood and biochemical test data for diagnosis of neurological diseases. *Sci Rep*. 2025;15(1):27857. Published 2025 Jul 30. doi:10.1038/s41598-025-09439-4.
8. Negrini D, Zecchin P, Ruzzenente A, et al. Machine Learning Model Comparison in the Screening of Cholangiocarcinoma Using Plasma Bile Acids Profiles. *Diagnostics (Basel)*. 2020;10(8):551. Published 2020 Aug 2. doi:10.3390/diagnostics10080551.
9. Farrell CJ, Makuni C, Keenan A, Maeder E, Davies G, Giannoutsos J. A Machine Learning Model for the Routine Detection of "Wrong Blood in Complete Blood Count Tube" Errors. *Clin Chem*. 2023;69(9):1031-1037. doi:10.1093/clinchem/hvad100.
10. Spies NC, Hubler Z, Azimi V, et al. Automating the Detection of IV Fluid Contamination Using Unsupervised Machine Learning. *Clin Chem*. 2024;70(2):444-452. doi:10.1093/clinchem/hvad207.
11. Auerbach M, Adamson JW. How we diagnose and treat iron deficiency anemia. *Am J Hematol*. 2016 Jan;91(1):31-8. doi: 10.1002/ajh.24201. Epub 2015 Nov 17.
12. Cook JD. Clinical evaluation of iron deficiency. *Semin Hematol*. 1982 Jan;19(1):6-18. PMID: 6763340.
13. Kuhn M. Building Predictive Models in R Using the caret Package. *J Stat Soft*. 2008;28(5):1-26. doi: 10.18637/jss.v028.i05.
14. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011 Mar 17;12:77. doi: 10.1186/1471-2105-12-77.
15. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988 Sep;44(3):837-45.
16. Rashidi HH, Tran N, Albahra S, Dang LT. Machine learning in health care and laboratory medicine: General overview of supervised learning and Auto-ML. *Int J Lab Hematol*. 2021;43 Suppl 1:15-22. doi:10.1111/ijlh.13537.
17. Tepakhan W, Srisintorn W, Penglong T, Saelue P. Machine learning approach for differentiating iron deficiency anemia and thalassemia using random forest and gradient boosting algorithms. *Sci Rep*. 2025;15(1):16917. Published 2025 May 15. doi:10.1038/s41598-025-01458-5.
18. Barton JC, Barton JC, Acton RT. Factors related to mean corpuscular volume in HFE p.C282Y homozygotes. *EJHaem*. 2024;6(1):e1063.
19. Garduno-Rapp NE, Ng YS, Weon JL, Saleh SN, Lehmann CU, Tian C, Quinn A. Early identification of patients at risk for iron-deficiency anemia using deep learning techniques. *Am J Clin Pathol*. 2024 Sep 3;162(3):243-251. doi: 10.1093/ajcp/aae031.

20. Efros O, Soffer S, Mudrik A, et al. Predictive machine-learning model for screening iron deficiency without anaemia: a retrospective cohort study. *BMJ Open*. 2025;15(8):e097016. Published 2025 Aug 12. doi:10.1136/bmjopen-2024-097016.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.