

Supplementary Table 6. Results Based on Evaluation in CRISP-DM (N = 125)

Index	Author (Year)	Classification and Regression Model Performance Metrics	Best model	Analysis of Feature Importance or Impact on Target Variable
A1	Marketta Hiissa et al. (2006)	AUC-ROC	Only one used (LS-SVM)	NI
A2	Manabu Nii et al. (2007)	Accuracy	Only one used (SVM-Sigmoid kernel)	NI
A3	Laurence G. Moseley et al. (2008)	Sensitivity, Specificity, Accuracy	Only one used (DT-CHAID)	NI
A4	Alexander Zlotnik et al. (2015)	MAE	M5P tree	NI
A5	Manabu Nii et al. (2016)	Accuracy	Only one used (SVM)	NI
A6	Robert Sherwin et al. (2017)	Sensitivity, Specificity, PPV	NI	NI
A7	Shinichiroh Yokota et al. (2017)	Sensitivity, Specificity.	Only one used (SVM-RBF)	NI
A8	Jeungok Choi et al. (2018)	Prediction rate	Only one used (LR)	NI
A9	Stephen I. Gallant et al. (2018)	AUC-ROC	LR (L2)	NI
A10	Eliezer Bose et al. (2019)	Accuracy	Only one used (GLM-Elastic Net)	Feature importance
A11	Gerald C. Gannod et al. (2019)	Recall, Precision, Accuracy, F1-score	Apriori+LR	NI
A12	Steven G. Johnson et al. (2018)	F2-score, Precision, Recall	SVM+Topic modeling	NI
A13	Zfania Tom Korach et al. (2019)	Concordance	Only one used (LDA)	NI
A14	Jae Yung Kwon et al. (2019)	AUC-ROC	XGBoost	Coefficients for LR, Relative gain improvement
A15	Suzanne S. Sullivan et al. (2019)	AUC-ROC, Sensitivity, Specificity	ANN	NI
A16	Maxim Topaz et al. (2019)	Precision, Recall, F-measure	Only one used (RF)	NI
A17	Heather Brom et al. (2020)	C-statistic	LR	Decision tree-based hierarchical branching
A18	Roschelle L. Fritz et al. (2020)	Accuracy, Sensitivity, Specificity, AUC-ROC	RF	NI
A19	Christopher M. Horvat et al. (2020)	AU-PRC, C-Statistic	RF	Boruta feature selection
A20	Renjie Hu et al. (2020)	R ²	ELMs+SOMs	NI
A21	Mireia Ladios-Martin et al. (2020)	Sensitivity, Specificity, AUC-ROC, Accuracy	Only one used (LR)	Permutation feature importance
A22	Soo-Kyoung Lee et al. (2020)	Accuracy, Sensitivity, Specificity, PPV, NPV	RF	Feature importance
A23	Chen Liang et al. (2020)	F1-score, Precision, Recall	NB	NI
A24	David S. Lindberg et al. (2020)	Sensitivity, Specificity, AUC-ROC	RF	Permutation importance, Variable ranking
A25	Jung In Park et al. (2020)	Accuracy, Sensitivity, Specificity, Precision, AUC-ROC	DT	Gini index
A26	Maxim Topaz et al. (2020)	Precision, Recall, F-score, AUC-ROC	RF	Information gain
A27	Dana M. Womack et al. (2020)	Accuracy, F1-score, AUC-ROC	GBM	RFE
A28	Ran An et al. (2021)	R2X(cum), R2Y(cum), Q2(cum)	PLS Regression	Feature importance
A29	Linyan Chen et al. (2021)	Accuracy, Sensitivity, Specificity, AUC-ROC	Only one used (SVM)	NI
A30	Aaron Conway et al. (2021)	AUC-ROC	RF	NI

(Continued on the next page)

Supplementary Table 5. Continued

Index	Author (Year)	Classification and Regression Model Performance Metrics	Best model	Analysis of Feature Importance or Impact on Target Variable
A31	Alberto Garcés-Jiménez et al. (2021)	Accuracy	RF	Random tree technique
A32	Li Hannaford et al. (2021)	Accuracy, AUC-ROC, Sensitivity, Specificity.	Stacked Model (NI)	Feature importance
A33	Farinaz Havaei et al. (2021)	R ² , RMSE	RF	importance score
A34	Elizabeth P. Howard et al. (2021)	NI	Only one used (Multiple LR)	NI
A35	Mingyue Hu et al. (2021)	Accuracy, Sensitivity, Specificity, AUC-ROC	NB	RFE
A36	Oleksandr Ivanov et al. (2021)	Accuracy, Sensitivity, Precision, F1-score, AUC-ROC	Only one used (XGBoost)	Information gain
A37	Liuqi Jin et al. (2021)	F1-score	RF	Gini index
A38	Jisu Kim et al. (2021)	Sensitivity, Specificity, Accuracy, AUC-ROC	Only one used (RF)	Mean decrease in accuracy
A39	Soo-Kyoung Lee et al. (2021)	Accuracy, Sensitivity, Specificity, PPV, NPV	RF	Importance score
A40	Chia-Hui Liu et al. (2021)	Accuracy, Sensitivity, Specificity, AUC-ROC	Bagging+RF	Information gain
A41	Tamara G. R. Macieira et al. (2021)	Accuracy, Precision, Recall, F-measure	Only one used (RF)	NI
A42	Takuro Nagata et al. (2021)	Accuracy, Precision, Jaccard index, Cohen's Kappa	SVM (RBF kernel)	NI
A43	Gojiro Nakagami et al. (2021)	Sensitivity, Specificity, AUC-ROC, PPV, NPV	XGBoost	Gini impurity
A44	Wenyu Song et al. (2021)	AUC-ROC, Accuracy, Sensitivity, Specificity, F1-score	RF	LR coefficients
A45	Rumei Yang et al. (2021)	Accuracy, Sensitivity, Specificity, PPV, NPV, F1-score, AUC-ROC	LR	Gini impurity
A46	Huaqiong Zhou et al. (2021)	AUC-ROC, Sensitivity, Specificity	GBM	Feature importance
A47	Yanhong Dong et al. (2022)	AUC-ROC	GBM	SHAP
A48	Farinaz Havaei et al. (2022)	R ² , RMSE	Only one used (RF)	Importance score
A49	Tingting Hu et al. (2022)	Accuracy, Recall, Precision, F1-score, AUC-ROC	LR	LR coefficient
A50	Shuai Jin et al. (2022)	AUC-ROC, Sensitivity, Specificity, Accuracy, Balanced Accuracy, Brier score, Calibration Curves	LR	Mean decrease in accuracy, Feature importance
A51	Mireia Ladios-Martin et al. (2022)	Accuracy, Precision, Recall, F1-Score, AUC-ROC	Bayes Point Machine	Average variable weights
A52	Young Ji Lee et al. (2022)	Macro F1-Score	Only one used (LR)	Top word features listed
A53	Anup Kumar Mishra et al. (2022)	Sensitivity, Specificity, Accuracy, F1-score, AUC-ROC	SVM	SHAP
A54	Kyoung Ja Moon et al. (2022)	Macro-averaged F1-score, Weighted F1-score, Accuracy, Sensitivity, Specificity, AUC-ROC	LEM2	NI

(Continued on the next page)

Supplementary Table 5. Continued

Index	Author (Year)	Classification and Regression Model Performance Metrics	Best model	Analysis of Feature Importance or Impact on Target Variable
A55	Nikhil Padhye et al. (2022)	Accuracy, AUC-ROC	ANN	NI
A56	Dongni Qian et al. (2022)	Accuracy	RF	Feature importance
A57	Javier Rojo et al. (2022)	MAE	NI	NI
A58	Jiyoun Song et al. (2022)	F-score, PRC	RF, SVM	Feature selection
A59	Tobias R. Spiller et al. (2022)	AUC, Sensitivity, Specificity, PPV, NPV	GBM	Relative importance of variables
A60	Katie Walker et al. (2022)	MAE	RF, LR	Feature importance
A61	Melyana Nurul Widyawati et al. (2022)	Precision, Recall, F1-score, Accuracy	DT	NI
A62	Jie Xu et al. (2022)	AUC-ROC, Sensitivity, Specificity, Accuracy, Precision, F1-score, PPV, NPV	LR	Gini impurity, LR coefficients
A63	Olga Yakusheva et al. (2022)	Cohen's Kappa, AUC-ROC	Only one used (RF)	Feature importance
A64	Ayla İrem Aydın et al. (2023)	ICC	Only one used (LR)	NI
A65	Rui CHEN et al. (2023)	AUC-ROC, Accuracy	RF	NI
A66	Ya-Huei Chen et al. (2023)	Accuracy, AUC-ROC, Sensitivity, Specificity, F1-Score	XGBoost	NI
A67	Pei-Yu Dai et al. (2023)	Accuracy, AUC-ROC	RF	SHAP
A68	Odai Y. Dweekat et al. (2023)	Sensitivity, Specificity, AUC-ROC, FPR, Detection Prevalence	Only one used (SVM)	RFE, Gini impurity
A69	Juliet Edgcomb et al. (2023)	Sensitivity	RF, LR (L1)	NI
A70	Ajeet Gajra et al. (2023)	Reduction rates in ED visits and hospital admissions	NI	NI
A71	Farinaz Havaei et al. (2023)	RMSE, R ²	Only one used (RF)	Importance scores
A72	Sharon Hewner et al. (2023)	Elbow method, Domain knowledge	Only one used (K-means)	Nursing knowledge
A73	Sunho Im et al. (2023)	AUC-ROC, Accuracy, F1-score, Sensitivity, Specificity, PPV	XGBoost	Information gain
A74	Junglyun Kim et al. (2023)	AUC-ROC, Accuracy, Sensitivity, Specificity	LR	Gini impurity
A75	Seong-Kwang Kim et al. (2023)	Accuracy, Precision, Recall, F1-score, AUC-ROC	RF	Feature importance
A76	Hyungbok Lee et al. (2023)	Accuracy, Precision, Recall, F-measure, AUC-ROC	LR	LR coefficients
A77	Hyungbok Lee et al. (2023)	Accuracy, Sensitivity, Specificity, AUC-ROC, PPV, NPV	RF	Gini impurity
A78	Lin-Lin Lee et al. (2023)	Correlation coefficients, ICC, Kappa values	Only one used (K-means)	NI
A79	Lingjuan Li et al. (2023)	AUC-ROC, Accuracy, Specificity, Sensitivity	RF	Feature importance
A80	Pei-Hung Liao et al. (2023)	Accuracy	Only one used (RF)	NI
A81	Sarah R. Martha et al. (2023)	Accuracy	Only one used (RF)	Mean decrease in accuracy
A82	Aruna Jothi Shanmugam et al. (2023)	Accuracy, Precision, Recall, F-Score	SVM	NI

(Continued on the next page)

Supplementary Table 5. Continued

Index	Author (Year)	Classification and Regression Model Performance Metrics	Best model	Analysis of Feature Importance or Impact on Target Variable
A83	Araceli Rodríguez Vico et al. (2023)	Confidence	Apriori	Decision tree importance
A84	Zeping Yan et al. (2023)	Accuracy, Sensitivity, Specificity, Precision, F1-score, AUC-ROC, Brier score	RF	Feature importance
A85	Metin Yildiz et al. (2023)	RMSE, MAE	RF	SHAP
A86	Ying Zhou et al. (2023)	F1-score, Accuracy, Precision, Recall, AUC-ROC	RF	SHAP
A87	Maryam Zolnoori et al. (2023)	AUC-ROC, F1-score	SVM	JMIM
A88	Young-Taek Park et al. (2024)	R ² , MSE, MAE, RMSE	Inpatient: RF, Outpatient: LR	Permutation feature importance
A89	Cynthia ABI KHALIL et al. (2024)	AUC-ROC, F1-score	XGBoost	Feature importance
A90	Yaser ALQARRAIN et al. (2024)	NI	Only one used (K-means)	NI
A91	Komal Aryal et al. (2024)	AUC-ROC, F1-score, Accuracy	GBM	Permutation importance
A92	Ranjana Chavan et al. (2024)	AUC-ROC, Accuracy, F1-score, Precision, Recall, MCC	CN2 Rule Inducer	NI
A93	Xiaomei Chen et al. (2024)	Accuracy, Sensitivity, Specificity, AUC-ROC	DT	Feature importance
A94	Colum Crowe et al. (2024)	Accuracy	LR	SHAP
A95	Tian Dai et al. (2024)	AUC-ROC, Specificity, Sensitivity, Accuracy, PPV, NPV, F1-score.	RF	SHAP
A96	Martha Duarte et al. (2024)	AUC-ROC, Accuracy, Precision, Recall, F1-score.	RF	Feature relevance scores
A97	Yu-Fang Guo et al. (2024)	AUC-ROC, Accuracy, Sensitivity, Specificity, Brier score	GBM	Permutation importance
A98	Rui Jin et al. (2024)	Accuracy	LR (L1)	NI
A99	Arisa Kawashima et al. (2024)	Accuracy, Precision, Sensitivity, Specificity, F-score, AUC-ROC	Only one used (XGBoost)	Recursive feature elimination
A100	Yeonju Kim et al. (2024)	AUC-ROC	RF	Feature importance
A101	Ju Hee Lee et al. (2024)	AUC-ROC, Sensitivity, Specificity, Accuracy, PPV, NPV	GBM	Variable contribution
A102	Pin-Chieh Lee et al. (2024)	AUC-ROC	Bayesian network	Neural network importance
A103	Renee C. B. Manworren et al. (2024)	Accuracy, Precision, Recall, AUC-ROC	RF	Feature importance
A104	Ninon Girardon da Rosa et al. (2024)	Accuracy, Sensitivity, F1-score, AUC-ROC	Only one used (RF)	Feature importance
A105	Jihye Kim Scroggins et al. (2024)	AUC-ROC, F1-score, Precision, Recall	XGBoost	Boruta feature selection
A106	Lu Shao et al. (2024)	AUC-ROC AU-PRC, Sensitivity, Specificity, Accuracy, Balanced Accuracy, PPV, NPV, F1-score	GBM	SHAP

(Continued on the next page)

Supplementary Table 5. Continued

Index	Author (Year)	Classification and Regression Model Performance Metrics	Best model	Analysis of Feature Importance or Impact on Target Variable
A107	Madeleine Stanik et al. (2024)	AUC-ROC, Accuracy, Recall, Sensitivity, Specificity, PPV, NPV, Precision	RF, XGBoost	SHAP
A108	Imam Tahyudin et al. (2024)	Accuracy, Precision, Recall, F1-Score, AUC-ROC	XGBoost	Feature importance
A109	Metin Yıldız et al. (2024)	RMSE, MAE	RF	SHAP
A110	Cheng Yu et al. (2024)	NI	XGBoost	Information gain
A111	Wei Zhang et al. (2024)	F2-Score, Accuracy, AUC-ROC, RF Sensitivity, Specificity	RF	Feature importance
A112	Maryam Zolnoori et al. (2024)	AUC-ROC, F1-score	SVM-RBF	JMIM
A113	Yunping Zhang et al. (2022)	Elbow method	Only one used (K-means)	NI
A114	Goran Erfani et al. (2024)	Silhouette measure of cohesion and separation	Only one used (K-means)	NI
A115	Tae Youn Kim et al. (2006)	Sensitivity, specificity, PPV, NPV, AUC-ROC	LR	Regression coefficients
A116	In Sook Cho et al. (2011)	Sensitivity, Specificity, PPV, NPV, AUROC	Only one used (BN)	Mutual information for BN model
A117	Yoko Setoguchi et al. (2016)	Accuracy, Sensitivity, Specificity	Only one used (DT)	Information gain
A118	Mikyung Moon et al. (2017)	Accuracy, Sensitivity, Specificity	Only one used (DT)	Decision tree splits (Gini impurity or information gain)
A119	Pacharmon Kaewprag et al. (2017)	Sensitivity, Specificity, PPV, NPV, AUC-ROC	BN	Markov Blanket property
A120	Xiaohong Deng et al. (2017)	Sensitivity, specificity, PPV, NPV, AUC-ROC	DT	NI
A121	Hong-Lin Chen et al. (2018)	AUC-ROC	Only one used (ANN)	Normalized importance calculation
A122	Hsiu-Lan Li et al. (2019)	Accuracy, Sensitivity, Specificity	SVM	NI
A123	Seul Ki Park (2020)	Sensitivity, Specificity, PPV, NPV, AUC-ROC	Only one used (LR)	Odds ratio
A124	Sookyung Hyun et al. (2021)	AUC-ROC, sensitivity, specificity, PPV, NPV	Only one used (LR)	NI
A125	Ji-Yu CAI et al. (2021)	AUC-ROC, sensitivity, specificity, PPV, NPV	Only one used (XGBoost)	Feature importance

ANN: Artificial Neural Network; AU-PRC: Area Under the Precision-Recall Curve; AUC-ROC: Area Under the Receiver Operating Characteristic Curve; C-statistic: Concordance Statistic; CN2 Rule Inducer: A rule-based machine learning algorithm for classification; DT: Decision Tree; DT-CHAID: Decision Tree using CHAID (Chi-squared Automatic Interaction Detection); ELMs: Extreme Learning Machines; FPR: False Positive Rate; GBM: Gradient Boosting Machine; GLM-Elastic Net regularization: Generalized Linear Model with Elastic Net regularization; ICC: Intraclass Correlation Coefficient; JMIM: Joint Mutual Information Maximization; K-means: Clustering algorithm that partitions data into k clusters; LDA: Latent Dirichlet allocation; LEM2: Learning from Examples Module, version 2; LR: Logistic Regression; LR (L1): Logistic Regression with L1 regularization; LS-SVM: Least Squares Support Vector Machine; M5P tree: M5P regression tree (model tree algorithm for regression); MAE: Mean Absolute Error; MCC: Matthews Correlation Coefficient; MSE: Mean Squared Error; NB: Naive Bayes; NI: No Information; NPV: Negative Predictive Value; PLS Regression: Partial Least Squares Regression; PPV: Positive Predictive Value; PRC: Precision-Recall Curve; Q2(cum): Cumulative Predictive Ability; R2X(cum): Cumulative Explained Variance for X Variables; R2Y(cum): Cumulative Explained Variance for Y Variables; RF: Random Forest; RFE: Recursive Feature Elimination; SHAP: Shapley additive explanations; SOMs: Self-Organizing Maps; SVM: Support Vector Machine; SVM-RBF kernel: Support Vector Machine with Radial Basis Function kernel; XGBoost: Extreme Gradient Boosting.