

Article

Comparison of Feature Selection Methods—Modelling COPD Outcomes

Jorge Cabral ^{1,*}, Pedro Macedo ¹, Alda Marques ² and Vera Afreixo ¹

¹ Center for Research and Development in Mathematics and Applications (CIDMA), Department of Mathematics, University of Aveiro, 3810-193 Aveiro, Portugal; pmacedo@ua.pt (P.M.); vera@ua.pt (V.A.)

² Respiratory Research and Rehabilitation Laboratory (Lab3R), School of Health Sciences (ESSUA) and Institute of Biomedicine (iBiMED), University of Aveiro, 3810-193 Aveiro, Portugal; amarques@ua.pt

* Correspondence: jorgecabral@ua.pt

Abstract: Selecting features associated with patient-centered outcomes is of major relevance yet the importance given depends on the method. We aimed to compare stepwise selection, least absolute shrinkage and selection operator, random forest, Boruta, extreme gradient boosting and generalized maximum entropy estimation and suggest an aggregated evaluation. We also aimed to describe outcomes in people with chronic obstructive pulmonary disease (COPD). Data from 42 patients were collected at baseline and at 5 months. Acute exacerbations were the aggregated most important feature in predicting the difference in the handgrip muscle strength (dHMS) and the COVID-19 lockdown group had an increased dHMS of 3.08 kg (CI₉₅ ≈ [0.04, 6.11]). Pack-years achieved the highest importance in predicting the difference in the one-minute sit-to-stand test and no clinical change during lockdown was detected. Charlson comorbidity index was the most important feature in predicting the difference in the COPD assessment test (dCAT) and participants with severe values are expected to have a decreased dCAT of 6.51 points (CI₉₅ ≈ [2.52, 10.50]). Feature selection methods yield inconsistent results, particularly extreme gradient boosting and random forest with the remaining. Models with features ordered by median importance had a meaningful clinical interpretation. Lockdown seem to have had a negative impact in the upper-limb muscle strength.

Keywords: feature selection; stepwise selection; LASSO; Boruta; extreme gradient boosting; random forest; COPD

MSC: 62P10



Citation: Cabral, J.; Macedo, P.; Marques, A.; Afreixo, V. Comparison of Feature Selection Methods—Modelling COPD Outcomes. *Mathematics* **2024**, *12*, 1398. <https://doi.org/10.3390/math12091398>

Academic Editors: Miguel Ángel Montero-Alonso and Juan De Dios Luna del Castillo

Received: 28 March 2024

Revised: 21 April 2024

Accepted: 29 April 2024

Published: 3 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Regression models aim to describe and predict an outcome given the values of n -dimensional vectors of p input features [1,2]. The task can be challenging especially when $p \gg n$ [3,4]. It is possible that all of the features are associated with the outcome, but often only a subset of the collected features can be considered [5]. As p increases, trying out every possible subset of features can become unfeasible [5,6]. The problem of high-dimensional feature selection (FS), defined as the process of reducing dimensionality by removing irrelevant features and identifying the most important ones [7,8], has received tremendous attention during the last decades. Although FS can help in obtaining models with less correlated features, biases, and unwanted noise, studies have shown that some of them can be 100% accurate using only non-informative features [9–11]. For instance, automatic stepwise selection has well reported limitations, such as the sensitivity to the presence of nuisance features and collinearity [12] which are exacerbated in the context of big data and the intensive computing time required [5,13,14].

The definition of ‘importance’ is also controversial since it may depend on subjective criteria assumed by the user whatever the technique is considered. Algorithms such as random forest (RF) [15], Boruta [16] or extreme gradient boosting (XGB) [17] provide measures that sort features according to their importance despite enhancing accuracy at

the expense of interpretability [5,15]. RF is an ensemble learning technique that uses the predictions of a set of decision trees computed in a bootstrap sample with a random subset of features in order to produce an aggregated result. Boruta is an algorithm designed as a wrapper that extends data by creating shuffled copies of all features and then trains a RF classifier in order to iteratively remove features deemed highly unimportant based on a chosen feature importance measure and a computed Z score. XGB is a computationally efficient ensemble learning algorithm that iteratively uses decision trees as weak learners along with regularization and a gradient descent optimization technique in order to enhance generalization and prevent overfitting. Penalty techniques, which force some of the estimated coefficients to be equal or close to zero, e.g., the least absolute shrinkage and selection operator (LASSO) method, can also perform FS [18].

A different approach supported by the information theory and info-metrics can be used [19–21]. Normalized entropy (NE), based on the consistent and asymptotically normal generalized maximum entropy (GME) estimator [22], measures the information content of a particular model or feature [23] and therefore can be used for FS.

FS can be applied in multiple fields of knowledge. For instance, studies suggest that well trained models provide clinically meaningful features with precision [24]. Also, selecting features associated with patient-centered outcomes is extremely important because it can lead to personalized and effective treatments for several diseases [25,26].

Chronic obstructive pulmonary disease (COPD) is a progressive, treatable and preventable respiratory disease [27]. It is the third-leading cause of death worldwide, killing 3.2 million individuals every year and accounts for a substantial individual, economic and societal burden [28,29]. Morbidity and prevalence seem to increase with age [30,31]. Although cigarette smoking is the leading COPD environmental risk factor, sex, genetics and comorbidities also seem to play an important role on the disease development and progression [32]. Also, body mass index (BMI) is associated with the rate of lung function decline where obesity seems to be protective [33,34]. External factors, such as the 2020 imposed lockdown due to the coronavirus disease 2019 (COVID-19) pandemic, may also influence the disease trajectory. For instance, a significant reduction of acute exacerbations of COPD (AECOPD) and COPD-related emergency department attendances during the lockdown period was found [35–37]. Also, an improvement in symptoms, and a significant reduction in COPD-related health care costs occurred during this period. On the other hand, the severity of participant's dyspnea worsened [38]. Although a significant increase in body weight was found in the general population [39], patients with COPD tended to lose weight during lockdown [40].

Our objectives in this work were to compare the results of different FS methods, including the promising yet underexplored approach of normalized entropy, analyze the correlation between results of different FS methods, illustrate how misleading their individual interpretation can be, and suggest an aggregated evaluation for the results of FS methods. Additionally, we also aimed to describe the effect of the COVID-19 lockdown, sociodemographic and clinical features on the lower- and upper-limb functional status and impact of the disease in people with stable COPD.

2. Materials and Methods

This section describes the study, in particular, participants, data collected and statistical techniques employed.

2.1. Study Design and Participants

Data collected between January 2019 and July 2020 in GENIAL (PTDC/DTP-PIC/2284/2014) and PRIME (PTDC/SAU-SER/28806/2017) research projects were used. Individuals were eligible if diagnosed with COPD [27] and clinically stable over the previous month. Individuals with other respiratory diseases, signs of cognitive impairment or presence of a significant or unstable cardiovascular, neurological or musculoskeletal disease were excluded. Written informed consent was first obtained from all participants.

2.2. Data Collection

Sociodemographic, anthropometric and clinical data (e.g., Charlson comorbidity index (CCI) [41], use of long-term oxygen therapy (LTOT) and non-invasive ventilation (NIV)) were assessed with a structured questionnaire. Lung function (forced expiratory volume in one second (FEV₁) and the ratio between FEV₁ and the forced vital capacity (FVC)) was assessed with spirometry [42]. The modified British medical research council questionnaire (mMRC) [43,44], the modified Borg scale [45–47], the brief physical activity assessment tool (BPAAT) [48] and the Saint George's respiratory questionnaire (SGRQ) [49] were used.

Upper and lower-limb functional status were assessed with the handgrip muscle strength (HMS) [50] and one-minute sit-to-stand test (1minSTS) [51,52]. Minimal clinically important differences (MCID) of 5.0 kg [53] and three repetitions [54] were considered. The COPD assessment test (CAT) evaluated the disease impact of the disease [55,56] and an MCID of two-points was considered [57].

Data were collected cross-sectionally at baseline and assessments with the 1minSTS, HMS and CAT were repeated after five months (post).

2.3. Statistical Analysis

Data were split in two groups; participants with baseline date between the 1 February 2019 and the 15 March 2019 were classified as pre-lockdown and participants with baseline date between the 1 February 2020 and the 15 March 2020 were classified as lockdown.

Variables were summarized accordingly. Shapiro-Wilk test was used to assess the assumption of normality. Welch *t*-tests and Mann-Whitney-Wilcoxon tests were used to compare characteristics between groups. Cohen's *d* effect size, phi coefficient and Cramer's *V* were calculated to assess association between variables. Chi-squared tests with simulated *p*-values for small cell sizes were used to compare proportions of baseline characteristics between groups.

The difference (*d*) between baseline and post values of the HMS, 1minSTS and CAT was determined and modelled by applying seven algorithms on numeric standardized data: (i) LASSO; (ii) Akaike's information criterion (AIC) [58] based automatic stepwise selection (stepAIC); (iii) Bayesian information criterion (BIC) [59] based automatic stepwise selection (stepBIC), (iv) normalized entropy; (v) RF; (vi) Boruta; (vii) XGB.

A preliminary tuning of RF parameters was performed with a grid of values for the number of features to consider at each split point (*mtry*) and the minimum number of observations in a terminal node (*nodesize*). The pair of values that produced the lowest out-of-bag (OOB) error [60,61] was used in 1000 trees. Feature importance was determined based on how much the accuracy decreased when the feature was excluded, given in percentage of the mean squared error (MSE).

For the Boruta algorithm, variables were classified as confirmed important, unconfirmed and confirmed unimportant according with shadow features [16].

XGB models were trained in a 4-fold cross-validation process with 750 iterations using the values of a grid containing combinations of the learning rate (*eta*) = 0.010, 0.015, 0.020, 0.025, the subsampling = 0.4, 0.5, 0.6, the minimum child weight = 1, 2, 3 and the maximum depth of a tree = 5, 8, 10, 11, 12, 14, 17. A *gbtree* booster and an objective of *reg:squarederror* were used [62]. The iteration with the lowest root MSE (RMSE) was considered. Feature importance was defined by the fractional contribution of each feature to the model based on the total gain of this feature's splits [62].

The penalty parameter λ used in LASSO was the one that produced the lowest 5-fold cross-validation MSE from a grid of 15,000 log values ranging from -7 to -1 .

Automatic stepwise selection consisted of a backward elimination process from an ordinary least squares (OLS) linear model (LM) with all features in order to obtain the lowest AIC/BIC [63].

In the NE procedure [23,64], the definition of supports for the GME estimator was done according to [65], that is, the limits of each support are established by the absolute maximum values of the ridge estimates [66]. Has recently emerged an interest with this

approach, mainly because (1) it is simple to perform, (2) it allows the use of non-sample information, (3) it is free of asymptotic requirements, (4) it involves a shrinkage rule that reduces mean squared error, (5) it allows to account for model misspecifications and model uncertainty, and (6) it can be implemented for well- and ill-posed models, including ill-conditioned models and small sample sizes (micronumerosity).

Features were ordered by their median importance. In case of ties, the interquartile range was used. Kendall's rank coefficient of correlation (τ) was determined to measure the association between FS methods [67,68].

OLS LMs were applied to non-standardized data with an increasing number of ordered by median importance features. The model kept was the one with the best performance score calculated by normalizing AIC, BIC, coefficient of multiple determination (R^2), adjusted R^2 , RMSE and residual standard deviation (Sigma) and taking the three times repeated 5-fold cross-validation mean value for each model [69]. Assumptions were assessed by visual inspection of residuals. The assumption of homogeneity of variances was further validated with the Breusch-Pagan test. Estimated marginal means (predicted values) for specific model features were computed [70].

For the sake of simplicity, a significance level of 0.05 was considered, so that when $p < 0.05$ the corresponding null hypothesis is rejected.

Statistical analyses were performed using R packages JWileymisc [71], randomForestSRC [72], randomForest [73], Boruta [74], xgboost [62], glmnet [75] and MASS [76], performance [77], sjPlot [78] and ggeffects [70] in RStudio Version 2023.12.1+402 [79] running R version 4.3.3 [80].

3. Results

3.1. Descriptive Analysis

A total of 42 participants with COPD were included, 24 (57.1%) of whom belonging to the pre-lockdown group. Participants mean age was 66.3, with standard deviation of 7.8 years, most were men (81.0%), former smokers (85.7%) and presented 3 to 4 comorbidities (64.3%) (Table 1). No statistically significant differences between participants' characteristics of the pre-lockdown and the lockdown groups were found.

In the pre-lockdown group the difference of -1.95 kg between baseline and post HMS was statistically significant ($t(36) \approx -2.24$, $p \approx 0.036$) (Figure 1).

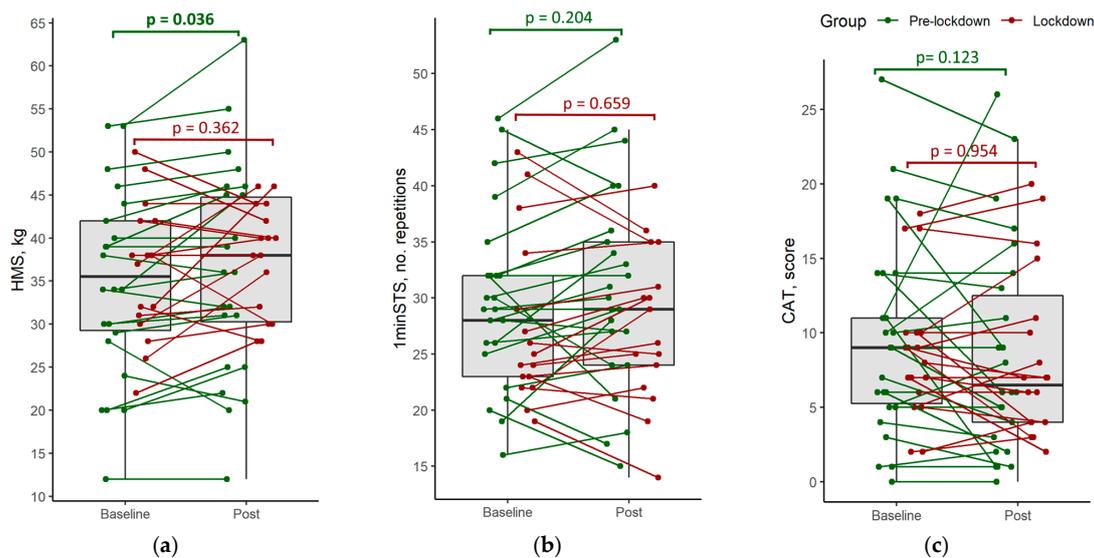


Figure 1. Distribution of participants' outcomes in the pre-lockdown ($n = 22$; 23; 24) and lockdown ($n = 16$; 16; 18) groups: (a) handgrip muscle strength (HMS); (b) number of repetitions in the one-minute sit-to-stand test (1minSTS); (c) points in the COPD assessment test (CAT). Note: p values (p) from Welch t -tests and Mann-Whitney-Wilcoxon tests.

Table 1. Descriptive statistics of participants’ characteristics at baseline ($n = 42$).

Characteristics		All ($n = 42$)	Pre-Lockdown ($n = 24$)	Lockdown ($n = 18$)	Tests
SEX					$\chi^2 \approx 0.12, p \approx 1.000, \varphi \approx 0.05$
	FEMALE	8 (19.0)	5 (20.8)	3 (16.7)	
	MALE	34 (81.0)	19 (79.2)	15 (83.3)	
AGE, YEARS		66.29 (7.83)	67.00 (7.97)	65.33 (7.75)	$t(40) \approx 0.68, p \approx 0.501, d \approx 0.21$
BODY MASS INDEX, KG/M ²		27.87 (5.24)	26.97 (5.71)	29.08 (4.42)	$t(40) \approx -1.30, p \approx 0.199, d \approx 0.41$
SMOKING STATUS					$\chi^2 \approx 2.64, p \approx 0.434, V \approx 0.25$
	NEVER	3 (7.1)	2 (8.3)	1 (5.6)	
	FORMER	36 (85.7)	19 (79.2)	17 (94.4)	
	CURRENT	3 (7.1)	3 (12.5)	0 (0.0)	
SMOKING NO. OF YEARS, YEARS		36.86 (15.40)	35.25 (15.91)	39.00 (14.86)	$t(40) \approx -0.78, p \approx 0.442, d \approx 0.24$
PACK-YEARS		63.03 (53.35)	64.12 (62.09)	61.57 (40.54)	$t(40) \approx 0.15, p \approx 0.880, d \approx 0.05$
CCI					$\chi^2 \approx 0.18, p \approx 1.000, V \approx 0.07$
	MILD (1–2)	9 (21.4)	5 (20.8)	4 (22.2)	
	MODERATE (3–4)	27 (64.3)	16 (66.7)	11 (61.1)	
	SEVERE (>=5)	6 (14.3)	3 (12.5)	3 (16.7)	
LTOT					$\chi^2 \approx 0.15, p \approx 1.000, \varphi \approx 0.06$
	NO	36 (85.7)	21 (87.5)	15 (83.3)	
	YES	6 (14.3)	3 (12.5)	3 (16.7)	
NIV					$\chi^2 \approx 0.27, p \approx 0.721, \varphi \approx 0.08$
	NO	32 (76.2)	19 (79.2)	13 (72.2)	
	YES	10 (23.8)	5 (20.8)	5 (27.8)	
AECOPD					$\chi^2 \approx 3.64, p \approx 0.189, V \approx 0.29$
	0	33 (78.6)	19 (79.2)	14 (77.8)	
	1	3 (7.1)	3 (12.5)	0 (0.0)	
	2 OR MORE	6 (14.3)	2 (8.3)	4 (22.2)	
RESP. EMERGENCIES					$\chi^2 \approx 0.00, p \approx 1.000, \varphi < 0.01$
	NO	35 (83.3)	20 (83.3)	15 (83.3)	
	YES	7 (16.7)	4 (16.7)	3 (16.7)	
RESP. HOSPITALIZATIONS					$\chi^2 \approx 0.12, p \approx 1.000, \varphi \approx 0.05$
	NO	39 (92.9)	22 (91.7)	17 (94.4)	
	YES	3 (7.1)	2 (8.3)	1 (5.6)	
FEV ₁ , % predicted		62.33 (23.31)	56.93 (24.25)	69.53 (20.48)	$t(40) \approx -1.78, p \approx 0.083, d \approx 0.55$
FEV ₁ /FVC, %		53.92 (12.06)	51.28 (12.91)	57.44 (10.13)	$t(40) \approx -1.67, p \approx 0.102, d \approx 0.52$
MMRC, points		1.26 (1.06)	1.42 (1.14)	1.06 (0.94)	$t(40) \approx 1.09, p \approx 0.280, d \approx 0.34$
BORG DYSPNOEA, points		0.80 (1.15)	0.60 (1.17)	1.06 (1.11)	$t(40) \approx -1.26, p \approx 0.213, d \approx 0.39$
BORG FATIGUE, points		1.10 (1.44)	1.00 (1.44)	1.22 (1.48)	$t(40) \approx -0.49, p \approx 0.627, d \approx 0.15$
BPAAT MODERATE, points		1.55 (1.56)	1.71 (1.55)	1.33 (1.61)	$t(40) \approx 0.76, p \approx 0.449, d \approx 0.24$
BPAAT VIGOROUS, points		0.14 (0.68)	0.25 (0.90)	0.00 (0.00)	$t(40) \approx 1.18, p \approx 0.245, d \approx 0.37$
SGRQ, points		32.79 (18.57)	36.64 (20.24)	27.66 (15.14)	$t(40) \approx 1.58, p \approx 0.122, d \approx 0.49$
HMS, KG, med [Q1, Q3] *					
	BASELINE	35.5 [29.3, 42.0]	34.0 [28.3, 41.5]	37.5 [30.8, 42.0]	$W \approx 163.0, p \approx 0.711$
	POST	38.0 [30.3, 44.8]	36.0 [26.5, 45.8]	39.0 [31.5, 42.5]	$W \approx 167.5, p \approx 0.813$
1MINSTS, no. rep., med [Q1, Q3] +					
	BASELINE	28.0 [23.0, 32.0]	29.0 [25.5, 32.0]	24.5 [22.8, 30.3]	$W \approx 225.5, p \approx 0.241$
	POST	29.0 [24.0, 35.0]	30.0 [25.5, 35.5]	27.5 [23.5, 32.0]	$W \approx 219.5, p \approx 0.317$
CAT, points, med [Q1, Q3]					
	BASELINE	9.0 [5.3, 11.0]	9.0 [5.0, 14.0]	8.5 [6.3, 10.0]	$W \approx 225, p \approx 0.828$
	post	6.5 [4.0, 12.5]	6.0 [2.8, 13.3]	7.0 [4.0, 10.8]	$W \approx 201.5, p \approx 0.721$

Note: Data presented as mean (standard deviation), count (percentage) or otherwise stated. Abbreviations: 1minSTS, one-minute sit-to-stand test; AECOPD, acute exacerbation of COPD; BPAAT, brief physical activity assessment tool; CAT, COPD assessment test; CCI, Charlson comorbidity index; COPD, chronic obstructive pulmonary disease; FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity; HMS, handgrip muscle strength; LTOT, long-term oxygen therapy; mMRC, modified medical council dyspnoea scale; no., number; NIV, non-invasive ventilation; SGRQ, St. George’s respiratory questionnaire; rep., repetitions; Resp., respiratory; med, median; Q, quartile; t, welch *t*-test statistics; *p*, *p*-value; *d*, Cohen’s *d*; *W*, Mann-Whitney-Wilcoxon statistics; χ^2 , Chi-squared statistics. * $n = 38$ (22/16); + $n = 39$ (23/16).

3.2. Handgrip Muscle Strength

3.2.1. Feature Importance

BORG fatigue score (4.9%) was considered the most important feature followed by AECOPD (4.0%) using the RF approach with an OOB error of 0.942 (Figures A1a and A2a). Boruta algorithm found the same two most important features but AECOPD (5.7) was confirmed important, while BORG fatigue score (5.0) was classified as unconfirmed (Figure A3a). FEV₁% predicted (0.16) was considered the most important feature by

the XGB algorithm (Table A1; Figure A4a). AECOPD was again the most important feature using LASSO with $\lambda \approx 1.45$ (Figure A5a,b). The AIC and BIC algorithm removed the same 13 features starting with CCI. With decreasing order of importance AECOPD, respiratory hospitalizations, FEV₁% predicted, age, BPAAT moderate score, sex, group and NIV were kept. AECOPD was the most important feature with a normalized entropy of 0.886 (Figure A6a) and was also the median most important feature (Figure 2a).

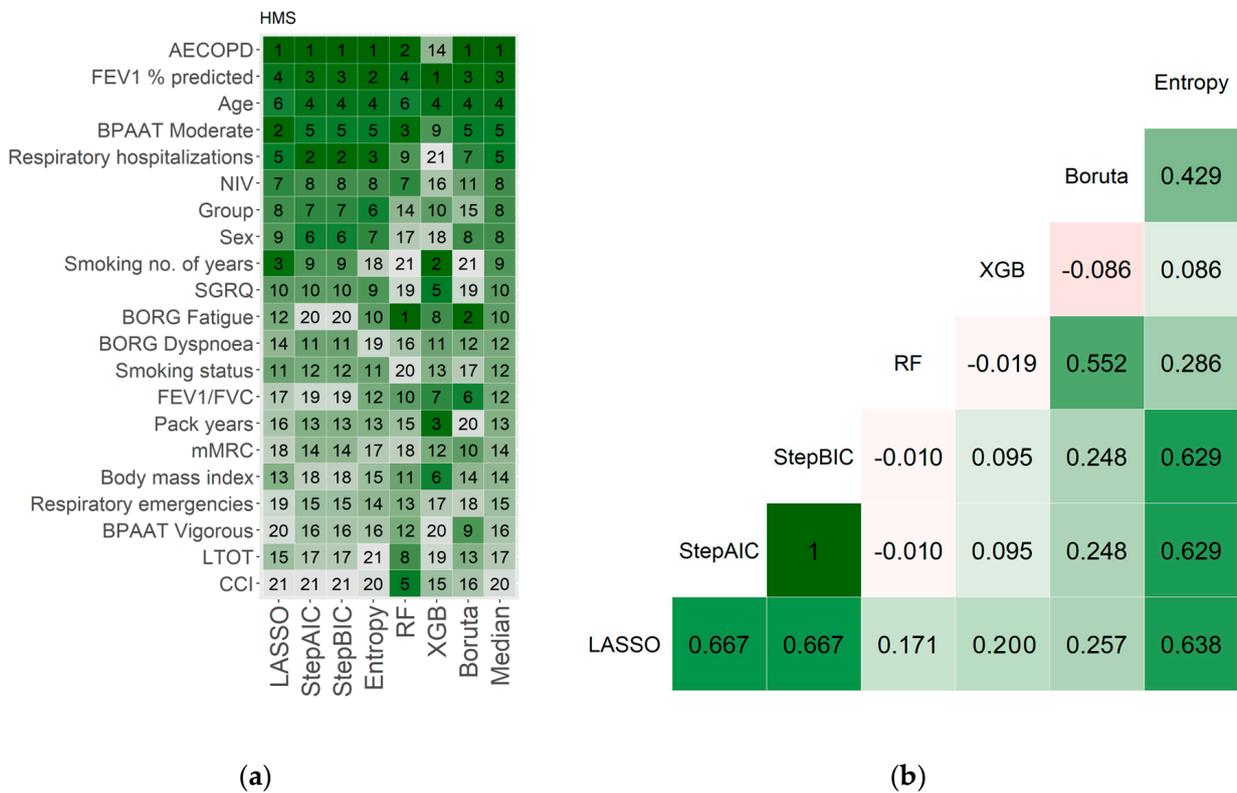


Figure 2. (a) Handgrip muscle strength (HMS) feature’s importance according to LASSO, AIC based stepwise automatic selection (StepAIC), BIC based stepwise automatic selection (StepBIC), normalized entropy (Entropy), random forest (RF), extreme gradient boosting (XGB) and Boruta algorithms in people with chronic obstructive pulmonary disease (COPD) ($n = 38$). The dark green to white gradient represent the decreasing of the features’ importance. (b) Kendall’s rank coefficient of correlation. The dark green to dark red gradient represent the decreasing of the value of Kendall’s rank coefficient of correlation, with white corresponding to zero. Abbreviations: AECOPD, acute exacerbation of COPD; BPAAT, brief physical activity assessment tool; BMI, body mass index; CCI, Charlson comorbidity index; FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity; LTOT, long-term oxygen therapy; mMRC, modified medical council dyspnoea scale; NIV, non-invasive ventilation; SGRQ, St. George’s respiratory questionnaire.

The stepwise methods agreed perfectly ($\tau = 1$), and the pairwise correlation between both stepwise methods and LASSO was high ($\tau \approx 0.676$) as it was between the entropy approach and LASSO ($\tau \approx 0.638$) (Figure 2b).

3.2.2. Linear Model

The LM generated with 8 features achieved the highest performance score (0.623) (Table 2). The residual analysis is available in Figure A7.

Under certain circumstances, participants with two or more AECOPD tend to improve their upper-limb strength more than the other participants. For instance, they are expected to have, on average, a decreased dHMS by 11.12 kg when compared with participants with no AECOPD ($CI_{95} \approx [6.36, 15.87]$; CI_{95} is the 95% Confidence Interval), ceteris paribus (everything else remains constant). Participants with respiratory hospitalizations tend to

have, on average, an increased dHMS by 7.32 kg (CI95 \approx [0.88, 13.76]), ceteris paribus. Every year added to a participant’s age results, on average, in an increase of 0.26 kg (CI95 \approx [0.03, 0.49]) in the dHMS, ceteris paribus. Finally, belonging to the lockdown group resulted, on average, in an increased dHMS by 3.08 kg (CI95 \approx [0.04, 6.11]), ceteris paribus (Table 2).

Table 2. Linear model’s coefficients and *p*-values for the handgrip muscle strength difference in people with chronic obstructive pulmonary disease using cumulatively the features ordered by median importance (*n* = 38).

	1 Feat	2 Feat	3 Feat	4 Feat	5 Feat	6 Feat	7 Feat	8 Feat
(Intercept)	−0.87	2.14	−7.58	−3.93	−5.71	−4.02	−5.17	−7.45
AECOPD [1]	−2.63	−2.07	−0.65	−0.21	−3.89	−3.25	−1.33	−1.41
AECOPD [>1]	−5.73 *	−6.68 **	−6.85 **	−7.30 **	−9.85 ***	−10.08 ***	−10.97 ***	−11.12 ***
FEV ₁ % predicted		−0.05	−0.05	−0.06	−0.04	−0.05	−0.08 *	−0.10 **
Age			0.15	0.13	0.14	0.13	0.16	0.26 *
BPAAT Moderate				−1.05 *	−1.10 *	−1.06 *	−1.04 *	−0.91 *
Hospitalizations [Yes]					7.21 *	6.98 *	6.69	7.32 *
NIV [Yes]						−2.07	−2.74	−3.07
Group [Lockdown]							2.63	3.08 *
Sex [Male]								−4.04
AIC	30.695	35.682	37.754	37.569	41.071	42.336	43.857	41.049
BIC	30.773	35.760	37.832	37.648	41.149	42.414	43.926	41.127
R ²	0.215	0.212	0.160	0.210	0.313	0.279	0.256	0.417
R ² adjusted	0.076	0.069	0.008	0.067	0.193	0.146	0.120	0.312
RMSE	4.860	4.934	5.005	4.631	4.257	4.827	4.827	4.091
Sigma	1.667	2.248	2.490	2.387	2.863	3.467	3.428	2.906
Performance score	0.599	0.400	0.245	0.392	0.463	0.234	0.159	0.623

Abbreviations: AECOPD, acute exacerbation of COPD; BPAAT, brief physical activity assessment tool; COPD, chronic obstructive pulmonary disease; feat, features; FEV₁, forced expiratory volume in 1 s; NIV, non-invasive ventilation; * *p* < 0.05; ** *p* < 0.01, *** *p* < 0.001.

Participants without hospitalizations and with two or more AECOPD tended to recover above the MCID. Generally, participants with respiratory hospitalizations in the previous year, with less than two AECOPD and caught in the lockdown tend to worsen above the MCID (Figure 3).

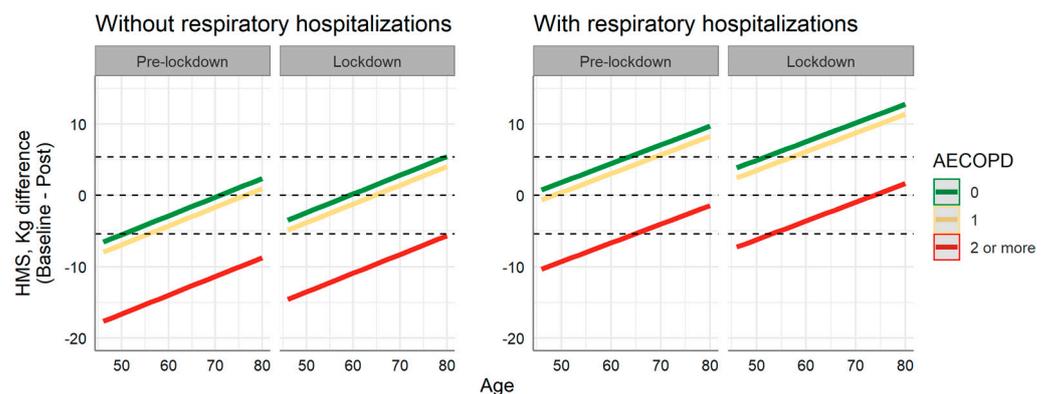


Figure 3. Predicted difference between baseline and post values in the handgrip muscle strength (HMS) of people with chronic obstructive pulmonary disease (COPD). Abbreviations: AECOPD, number of acute exacerbations of COPD. Note: predictions were made for male participants without non-invasive ventilation, with a brief physical activity assessment tool score of 0 and 70% of the predicted forced expiratory volume in 1 s. Dashed lines represent the minimal clinically important difference.

3.3. One-Minute Sit-to-Stand Test

3.3.1. Feature Importance

Pack-years (12.7%) had the highest importance value in the tuned RF algorithm (Figures A1b and A2b). Boruta algorithm found two confirmed important features, pack-years (7.2) and SGRQ (4.8), while sex (3.4) was classified as unconfirmed (Figure A3b). At 61 testing iterations (Table A1) the XGB algorithm also considered pack-years (0.24) the most important feature (Figure A4b).

LASSO with a penalty parameter of $\lambda \approx 1.34$ minimized the MSE and selected BORG Dyspnoea, sex and pack-years (Figure A5c,d). The AIC algorithm kept sex, BORG Dyspnoea, pack-years, SGRQ, mMRC, smoking status and FEV₁/FVC. Using BIC, BORG Dyspnoea and pack-years remained. Sex had the lowest normalized entropy (0.955) followed by pack-years (0.968) (Figure A6b). Pack-years achieved the highest median importance position (Figure 4a).

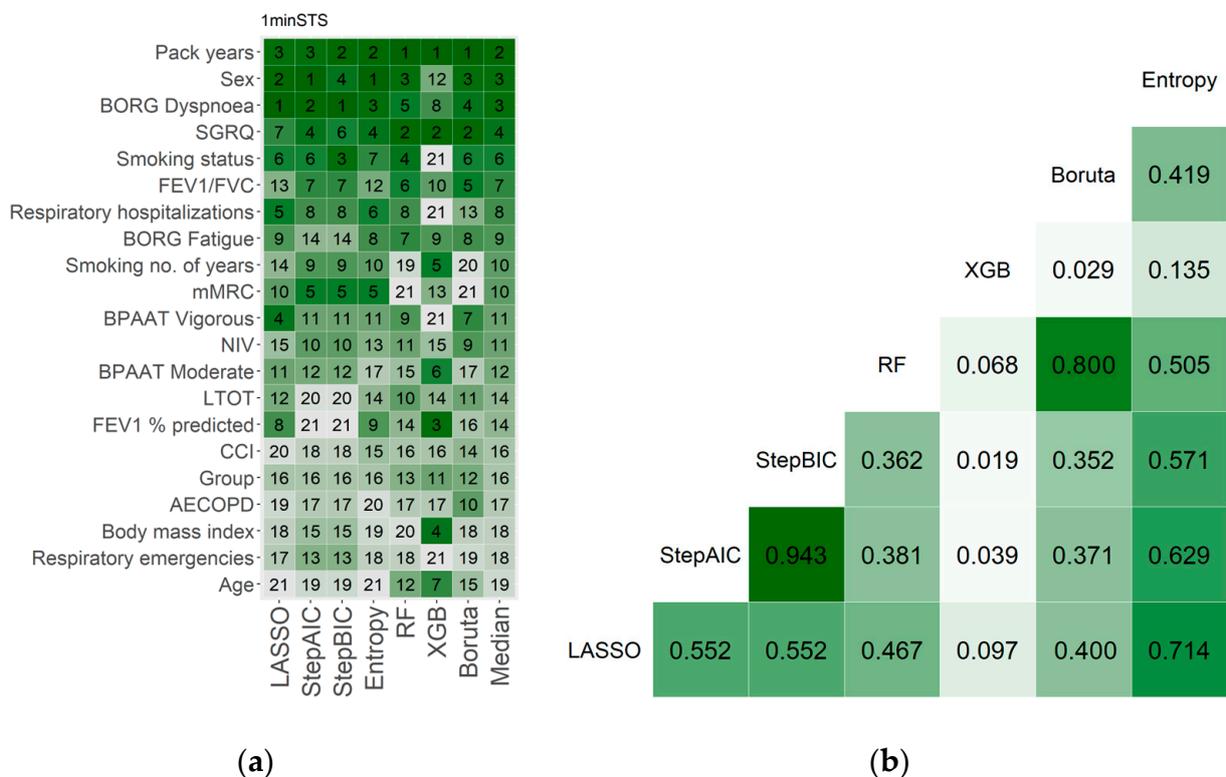


Figure 4. (a) One-minute sit-to-stand (1minSTS) feature’s importance according to LASSO, AIC based stepwise automatic selection (StepAIC), BIC based stepwise automatic selection (StepBIC), normalized entropy (Entropy), random forest (RF), extreme gradient boosting (XGB) and Boruta algorithms in people with chronic obstructive pulmonary disease (COPD) ($n = 39$). The dark green to white gradient represent the decreasing of the features’ importance. (b) Kendall’s rank coefficient of correlation. The dark green to dark red gradient represent the decreasing of the value of Kendall’s rank coefficient of correlation, with white corresponding to zero. Abbreviations: AECOPD, acute exacerbation of COPD; BPAAT, brief physical activity assessment tool; BMI, body mass index; CCI, Charlson comorbidity index; FEV1, forced expiratory volume in 1 s; FVC, forced vital capacity; LTOT, long-term oxygen therapy; mMRC, modified medical council dyspnoea scale; NIV, non-invasive ventilation; SGRQ, St. George’s respiratory questionnaire.

A high positive correlation was found between both stepwise methods ($\tau \approx 0.943$), and between Boruta and RF ($\tau \approx 0.800$). XGB returned correlation values approximately equal to zero with all other FS methods. The correlation between the entropy approach and LASSO was again high ($\tau \approx 0.714$) (Figure 4b).

3.3.2. Linear Model

The LM using the feature with highest median importance (residual analysis in Figure A8) had the highest performance score (0.951) (Table 3).

Table 3. Linear model’s coefficients for the difference in the number of repetitions of the one-minute sit-to-stand test in people with chronic obstructive pulmonary disease using cumulatively the features ordered by median importance ($n = 39$).

	1 Feat	2 Feat	3 Feat	4 Feat	5 Feat	6 Feat	7 Feat	8 Feat
(Intercept)	−2.75 *	−4.82 **	−5.24 **	−6.02 **	−4.67	−10.11	−9.79	−10.90
Pack-years	0.03 *	0.03	0.02	0.02	0.02	0.02	0.02	0.02
Sex [Male]	-	3.17	2.79	2.94	2.56	3.45	3.35	3.62
BORG Dyspnoea	-	-	1.26 *	1.09	1.10	1.09	1.02	0.51
SGRQ	-	-	-	0.03	0.02	0.04	0.04	0.03
Smoking status [Former]	-	-	-	-	−0.76	−0.29	−0.57	−0.49
Smoking status [Actual]	-	-	-	-	−4.64	−3.99	−4.15	−4.26
FEV1/FVC	-	-	-	-	-	0.07	0.07	0.09
Hospitalizations [Yes]	-	-	-	-	-	-	1.84	2.41
BORG Fatigue	-	-	-	-	-	-	-	0.56
AIC	27.227	32.500	37.007	39.504	43.710	40.896	42.837	40.827
BIC	27.376	32.649	37.156	39.641	43.837	41.032	42.964	40.976
R ²	0.465	0.273	0.236	0.149	0.211	0.231	0.235	0.060
R ² adjusted	0.376	0.147	0.099	−0.002	0.068	0.101	0.093	−0.105
RMSE	4.378	4.254	4.463	4.201	5.171	4.280	4.641	4.918
Sigma	1.423	1.770	2.107	2.480	3.458	2.726	3.626	2.694
Performance score	0.951	0.678	0.720	0.388	0.135	0.399	0.237	0.166

Abbreviations: FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity; SGRQ, St. George’s respiratory questionnaire; * $p < 0.05$; ** $p < 0.01$.

Participants with higher tobacco load tend to have their number of 1minSTS repetitions reduced over the lockdown period. On average, an increase of approximately 28.8 unit in pack-years tends to increase d1minSTS by 1 repetition (CI95 \approx [0.07, 1.93]). Participants do not tend to recover nor reduce their number of repetitions above the MCID (Figure 5).

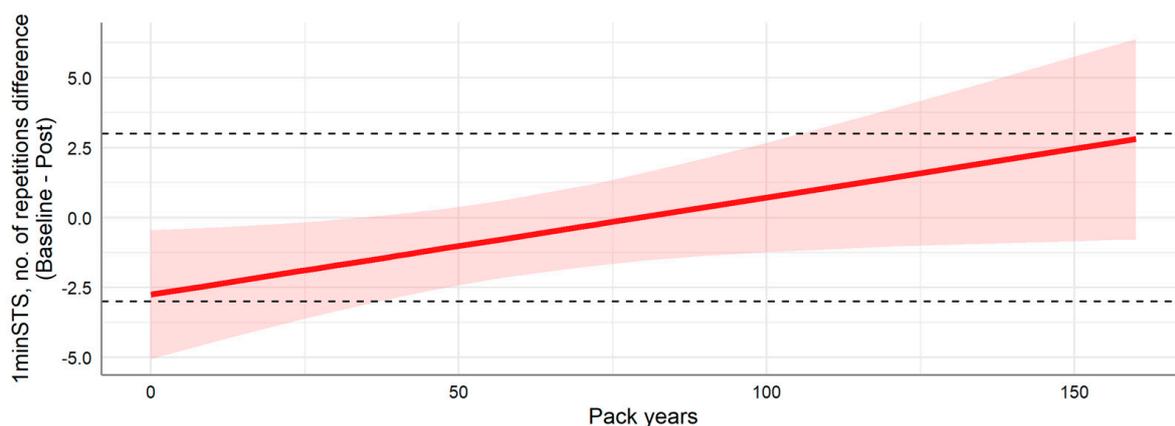


Figure 5. Predicted difference between baseline and post number of repetitions in the one-minute sit-to-stand test (1minSTS) of people with chronic obstructive pulmonary disease (COPD). Dashed lines represent the minimal clinically important difference.

3.4. COPD Assessment Test

3.4.1. Feature Importance

RF considered CCI (7.5%) the most important feature when mtry and nodesize were set at 2 and 13, respectively (Figures A1c and A2c). Boruta algorithm also confirmed as important CCI (6.5) and classified smoking no. of years (3.5) as unconfirmed (Figure A3c).

The lowest RMSE for the XGB algorithm was obtained for a learning rate eta of 0.020 and was achieved at 52 testing iterations (Table A1). Smoking no. of years (0.16) was considered the most important feature by XGB followed by SGRQ (0.13), pack-years (0.11) and age (0.10) (Figure A4c). CCI and existence of respiratory emergencies were selected by the BIC algorithm and LASSO with $\lambda \approx 1.26$ (Figure A5e,f). The AIC algorithm removed 18 features and kept CCI, AECOPD and SGRQ. CCI had the lowest normalized entropy (0.922) followed by the SGRQ (0.932) (Figure A6c). CCI had a median rank of 1 (Figure 6a).

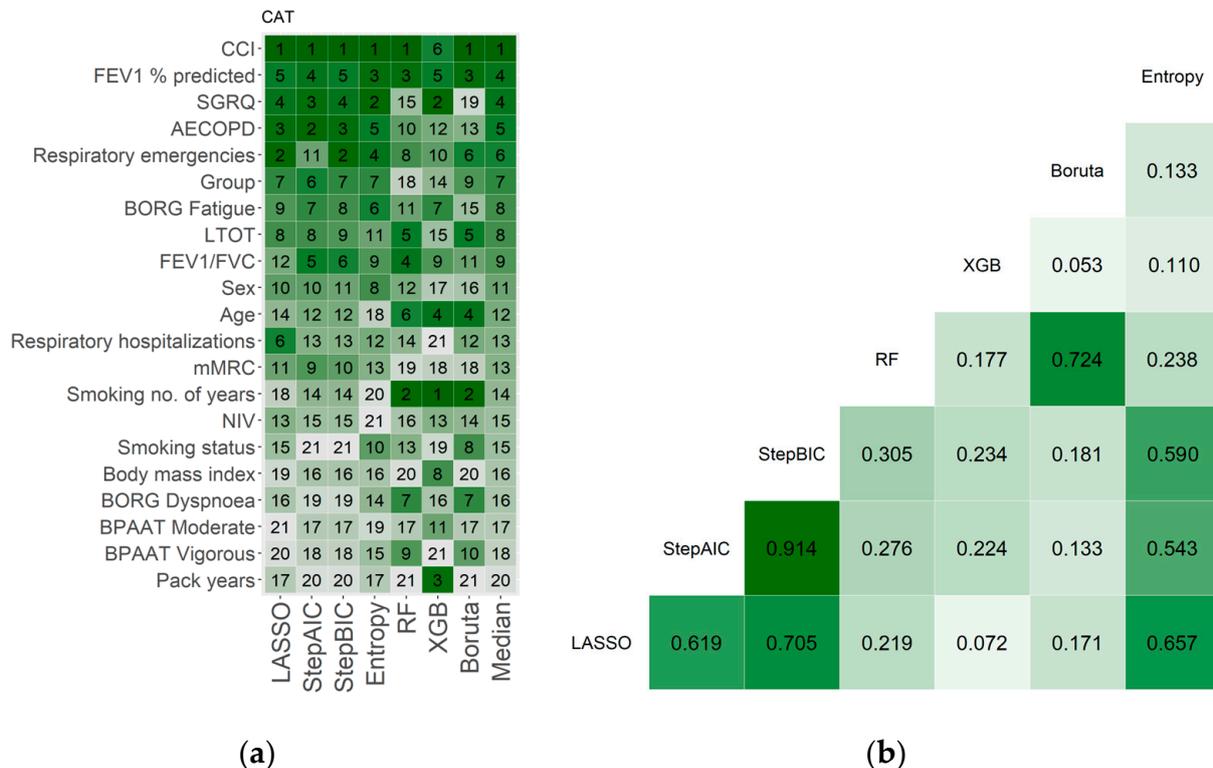


Figure 6. (a) COPD assessment test (CAT) feature’s importance according to LASSO, AIC based stepwise automatic selection (StepAIC), BIC based stepwise automatic selection (StepBIC), normalized entropy (Entropy), random forest (RF), extreme gradient boosting (XGB) and Boruta algorithms in people with chronic obstructive pulmonary disease (COPD) ($n = 42$). The dark green to white gradient represent the decreasing of the features’ importance. (b) Kendall’s rank coefficient of correlation. The dark green to dark red gradient represent the decreasing of the value of Kendall’s rank coefficient of correlation, with white corresponding to zero. Abbreviations: AECOPD, acute exacerbation of COPD; BPAAT, brief physical activity assessment tool; BMI, body mass index; CCI, Charlson comorbidity index; FEV1, forced expiratory volume in 1 s; FVC, forced vital capacity; LTOT, long-term oxygen therapy; mMRC, modified medical council dyspnoea scale; NIV, non-invasive ventilation; SGRQ, St. George’s respiratory questionnaire.

The pairwise correlation between both stepwise methods and LASSO was high, as it was between Boruta and RF ($\tau \approx 0.724$). The highest correlation with the entropy approach was obtained with LASSO ($\tau \approx 0.657$) (Figure 6b).

3.4.2. Linear Model

The highest performance score (0.859) was achieved by the LM with 4 features (Table 4, residual analysis in Figure A9).

Generally speaking, participants with severe CCI seem to have worsened their CAT score at the end of lockdown period. Specifically, participants with severe CCI are expected to have, on average, a decreased dCAT by 6.51 points when compared with participants with mild CCI ($CI_{95} \approx [2.52, 10.50]$), ceteris paribus. Those who have experienced one

AECOPD in the previous year are expected to have, on average, an increased dCAT by 4.97 points when compared with participants with no AECOPD (CI95 ≈ [0.09, 9.84]) and if at the same time, they have a mild CCI score they tend to recover above the MCID (Figure 7).

Table 4. Linear model’s coefficients for the difference in the COPD assessment test score in people with chronic obstructive pulmonary disease using cumulatively the features ordered by median importance ($n = 42$).

	1 Feat	2 Feat	3 Feat	4 Feat	5 Feat	6 Feat	7 Feat	8 Feat
(Intercept)	2.33	3.82	5.45	4.37	3.93	3.94	4.09	4.98
CCI [Moderate]	−1.07	−1.25	−1.56	−0.95	−0.88	−0.89	−0.86	−0.55
CCI [Severe]	−6.33 **	−6.45 **	−6.42 **	−6.51 **	−6.43 **	−6.43 **	−6.24 **	−5.97 **
FEV ₁ % predicted	-	−0.02	−0.03	−0.02	−0.03	−0.03	−0.03	−0.04
SGRQ	-	-	−0.03	−0.04	−0.03	−0.03	−0.04	−0.05
AECOPD [1]	-	-	-	4.97 *	4.37	4.36	4.66	4.95
AECOPD [>1]	-	-	-	2.44	0.88	0.89	0.50	1.03
Emergencies [Yes]	-	-	-	-	2.26	2.26	2.19	1.40
Group [Lockdown]	-	-	-	-	-	−0.01	−0.14	0.07
BORG Fatigue	-	-	-	-	-	-	0.45	0.53
LTOT [Yes]	-	-	-	-	-	-	-	−2.17
AIC	30.454	34.830	37.183	33.561	42.205	41.932	43.540	46.093
BIC	30.834	35.210	37.563	33.941	42.585	42.311	43.920	46.473
R ²	0.152	0.332	0.149	0.408	0.353	0.260	0.215	0.217
R ² adjusted	0.020	0.229	0.015	0.318	0.252	0.143	0.092	0.094
RMSE	4.060	4.288	4.104	4.221	4.194	4.417	4.393	4.577
Sigma	1.822	2.100	1.830	2.269	2.804	2.769	2.673	3.455
Performance score	0.671	0.707	0.508	0.836	0.534	0.352	0.278	0.087

Abbreviations: AECOPD, acute exacerbation of COPD; COPD, chronic obstructive pulmonary disease; CCI, Charlson comorbidity index; feat, features; FEV₁, forced expiratory volume in 1 s; LTOT, long-term oxygen therapy; SGRQ, St. George’s respiratory questionnaire; * $p < 0.05$; ** $p < 0.01$.

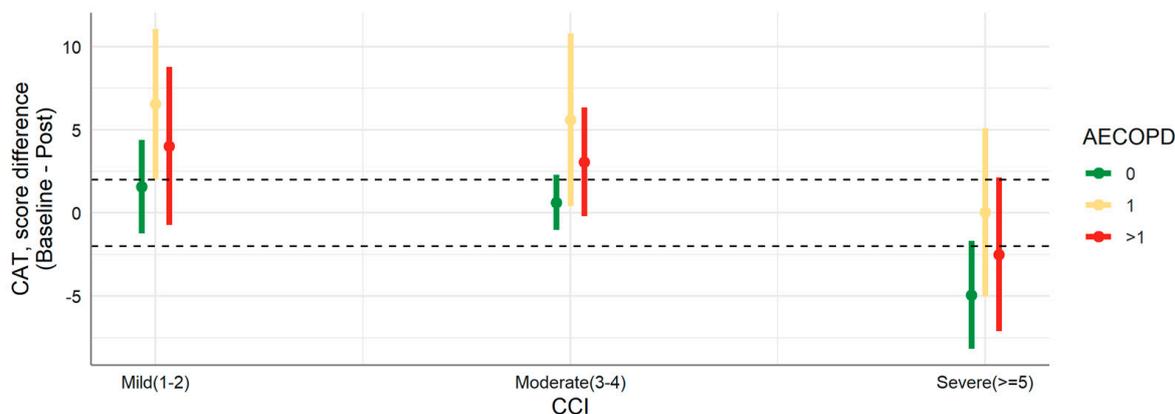


Figure 7. Predicted difference between baseline and post COPD assessment test (CAT) score of people with chronic obstructive pulmonary disease (COPD). Abbreviations: AECOPD, number of acute exacerbations of COPD; CCI, Charlson comorbidity index. Dashed lines represent the minimal clinically important difference.

4. Discussion

The main purpose of this study was to compare different common feature selection methods, including a rarely used one which is based on the normalized entropy, analyze the correlation between results of different FS methods and suggest an aggregated evaluation for the results, since the individual interpretation of FS methods can result in unreliable inferences [81,82]. Excessive number of features in health data is commonplace and FS is essential to simplify the prediction model’s learning process [81,83], so we also aimed to assess the relevance and clinical importance of the features selected when modelling

meaningful outcomes for people with COPD. Our study suggests that different FS methods attribute different importance to the same features. This finding seems to reinforce the uncertainty and heterogeneity associated to the selection of meaningful features also pointing out that there is no one-size-fits-all approach [6,84]. Different methodologies such as filter methods (e.g., association measures or test, information gain), wrapper methods (stepwise linear models, Boruta) or embedded methods (penalized regression models, extreme gradient boosting, random forest) are founded in different principles and have different theoretical structures. Therefore, the importance given to the same feature varies between them [84].

For instance, pack-years was considered the most important feature for predicting the difference in the number of repetitions of the 1minSTS and it was on the top 3 most important features for all FS methods. Yet, the number of AECOPD in the previous year, considered the most important feature for predicting dHMS, was the most important feature for five FS methods but XGB placed it at the 14th position. All but XGB considered CCI as the most important feature for predicting the dCAT, which ranked it at 6th position. For this outcome, AECOPD was on the top 3 of the most important features for LASSO and stepwise algorithms and studies found a significant association between the change in CAT scores and the risk of exacerbations [85]. If we only had considered the importance given by XGB, RF or Boruta this feature would not have been included in the final model. Also, the smoking number of years was considered the 1st or 2nd most important feature for RF, XGB and Boruta while the other four methods placed it between the 14th and the 20th position. In fact, XGB seems to be the FS method least associated with the remaining, although studies suggest that it produces models with improved accuracy, reduced misjudgment and great clinical significance [86,87]. For instance, although both are based on decision trees, RF and XGB may have produced different results given their different theoretical structure (aggregated solution vs. sequential solution). As expected, the automatic stepwise selection approaches produced similar results [88]. Studies found that Boruta could outperform either automatic selection or RF algorithms [89]. Our results showed a high correlation between the ranks of features produced by Boruta and RF algorithms. NE is consistently associated with LASSO.

Despite the existence of some COPD outcomes' prediction models, to our knowledge, none was obtained with data from individuals with COPD that were subjected to pulmonary rehabilitation before and immediately after the COVID-19 lockdown. The models obtained by our method suggest that the overall upper-limb muscle strength increase seems to be statistically smaller or the decrease tends to be statistically higher in the COVID-19 lockdown group. Having a higher comorbidity index seems to lead to a higher decline in the wellbeing of participants after five months. Nevertheless, participants with a lower index associated with respiratory emergencies perceive a recovery of their wellbeing after the same period of time. Aging and being hospitalized by respiratory causes have a negative effect on the evolution of the overall upper-limb muscle strength while a higher physical activity benefits its course. The study suggests that the follow-up performed by professionals, mainly by telephone, is an important strategy in order to prevent negative impact in the overall upper-limb muscle strength of patients with COPD, which is why it is advised when in-person monitoring is not available.

The strengths of our study include the comparison of different FS methods, one of them less commonly used although quite promising, and corresponding outcomes, which are interpreted by an aggregation procedure. Also, the use of real data gives the possibility to try to justify the relevance of selected features. Besides possible confounding factors that may occur [84], the study has some potential limitations: (1) real data with higher dimension of features and simulated data with different ratios between number of observations and number of features should be explored to assess the stability of the techniques, since there is evidence that they perform inconsistently [90]; (2) the pre-post design could be biased by seasonal trends; (3) mMRC and CAT were delivered face-to-face in the pre-lockdown period but telephonically during the lockdown. Yet, these are well known tools to both

participants and professionals; (4) the NE approach should be improved, considering, for instance, the generalized cross entropy estimator and the transformation group procedure usually adopted to construct priors in other contexts of maximum entropy estimation [20].

5. Conclusions

Feature selection methods can provide quite different results and should be used with caution. It is advisable not to be restrained to the use of only one method since the conclusions might be biased. Given previous clinical information, our linear models with features ordered by their median importance had a meaningful clinical interpretation. The generalization of the proposed median aggregation (an intuitive idea from robust statistics) to other contexts needs further scientific support through simulation studies. This study also showed that the restrictions to circulation, the social distancing and isolation resulting from COVID-19 pandemic seem to have had a negative impact in the overall upper-limb muscle strength of patients with COPD.

Author Contributions: Conceptualization, J.C.; methodology, J.C., P.M. and V.A.; formal analysis, J.C.; writing—original draft preparation, J.C.; writing—review and editing, J.C., P.M., A.M. and V.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Fundação para a Ciência e Tecnologia (FCT) through the Center for Research and Development in Mathematics and Applications (CIDMA) and projects <https://doi.org/10.54499/UIDB/04106/2020> and <https://doi.org/10.54499/UIDB/04106/2020>. This research was also supported by Programa Operacional de Competitividade e Internacionalização—POCI, through Fundo Europeu de Desenvolvimento Regional—FEDER (POCI01-0145-FEDER-028806), by Fundação para a Ciência e Tecnologia (PTDC/SAU-SER/28806/2017) and under the project UIDB/04501/2020.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy and ethical restrictions.

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

Appendix A

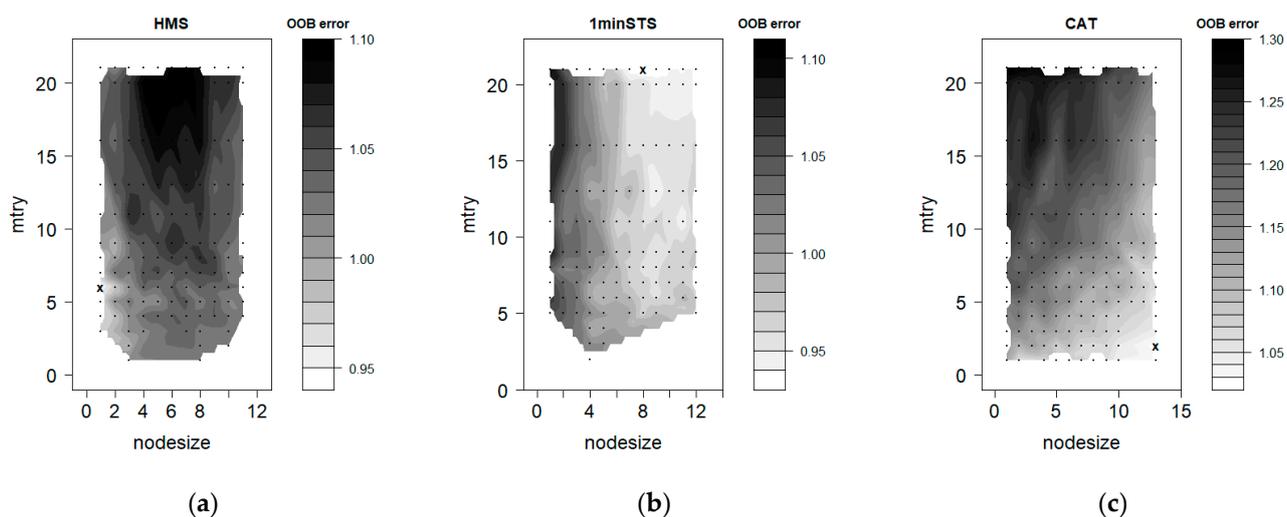


Figure A1. Random forest's out-of-bag (OOB) error for different values of number of features to consider at each split point (mtry) and minimum number of observations in a terminal node (nodesize). The parameters resulting in lowest OOB error are indicated with an x: (a) HMS, handgrip muscle strength; (b) 1minSTS, one-minute sit-to-stand test; (c) CAT, COPD assessment test.

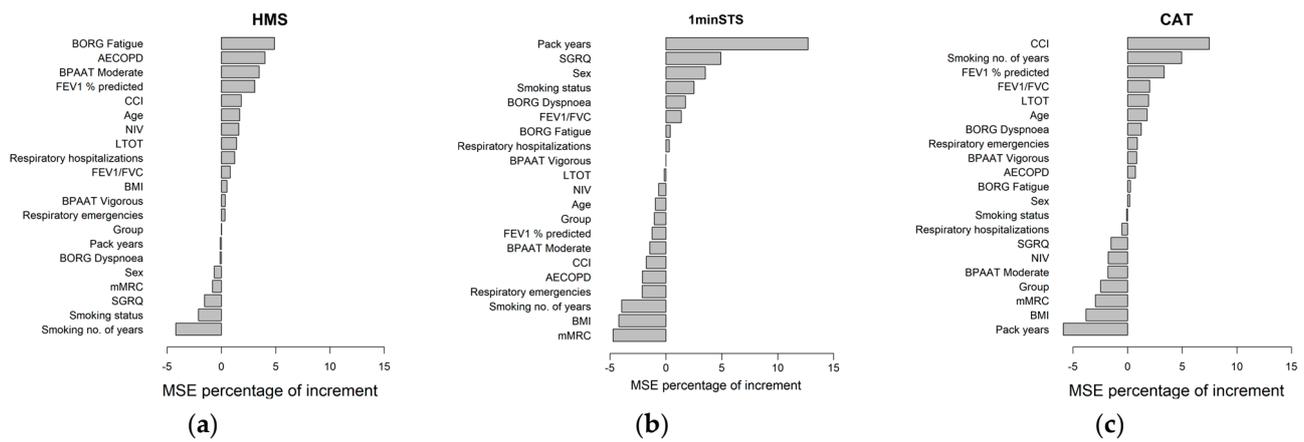


Figure A2. Feature importance given by the random forest algorithm for the difference in the outcomes in people with chronic obstructive pulmonary disease (COPD) ($n = 38; 39; 42$); (a) handgrip muscle strength (HMS); (b) one-minute sit-to-stand test (1minSTS); (c) COPD assessment test (CAT). Abbreviations: AECOPD, acute exacerbation of COPD; BPAAT, brief physical activity assessment tool; BMI, body mass index; CCI, Charlson comorbidity index; COPD, chronic obstructive pulmonary disease; FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity; LTOT, long-term oxygen therapy; MSE, mean squared error; mMRC, modified medical council dyspnoea scale; NIV, non-invasive ventilation; no., number; SGRQ, St. George’s respiratory questionnaire.

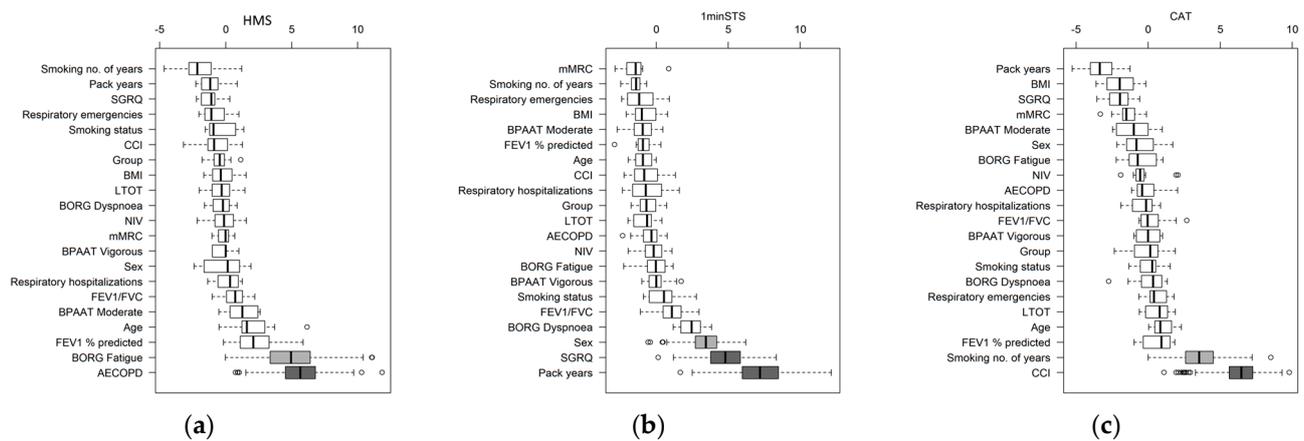


Figure A3. Feature importance given by the Boruta algorithm for the difference in the outcomes in people with chronic obstructive pulmonary disease (COPD) ($n = 38; 39; 42$); (a) handgrip muscle strength (HMS); (b) one-minute sit-to-stand test (1minSTS); (c) COPD assessment test (CAT). Dark grey corresponds to the confirmed important features, light grey corresponds to the unconfirmed features and white corresponds to confirmed unimportant features. Abbreviations: AECOPD, acute exacerbation of COPD; BPAAT, brief physical activity assessment tool; BMI, body mass index; CCI, Charlson comorbidity index; COPD, chronic obstructive pulmonary disease; FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity; LTOT, long-term oxygen therapy; mMRC, modified medical council dyspnoea scale; NIV, non-invasive ventilation; no., number; SGRQ, St. George’s respiratory questionnaire.

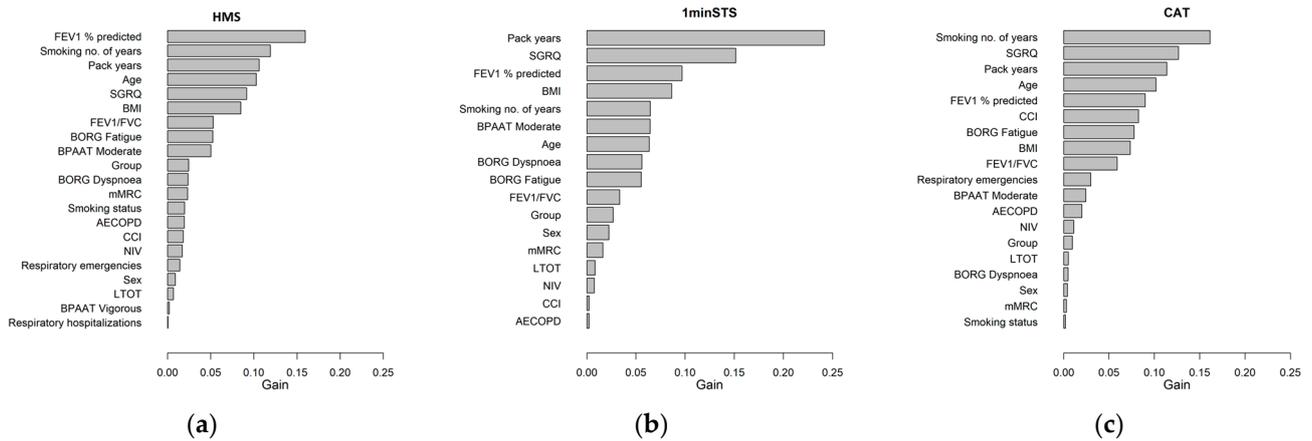


Figure A4. Feature importance given by the extreme gradient boosting algorithm for the difference in the outcomes in people with chronic obstructive pulmonary disease (COPD) ($n = 38; 39; 42$); (a) handgrip muscle strength (HMS); (b) one-minute sit-to-stand test (1minSTS); (c) COPD assessment test (CAT). Abbreviations: AECOPD, acute exacerbation of COPD; BPAAT, brief physical activity assessment tool; BMI, body mass index; CCI, Charlson comorbidity index; COPD, chronic obstructive pulmonary disease; FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity; LTOT, long-term oxygen therapy; mMRC, modified medical council dyspnoea scale; NIV, non-invasive ventilation; no., number; SGRQ, St. George’s respiratory questionnaire.

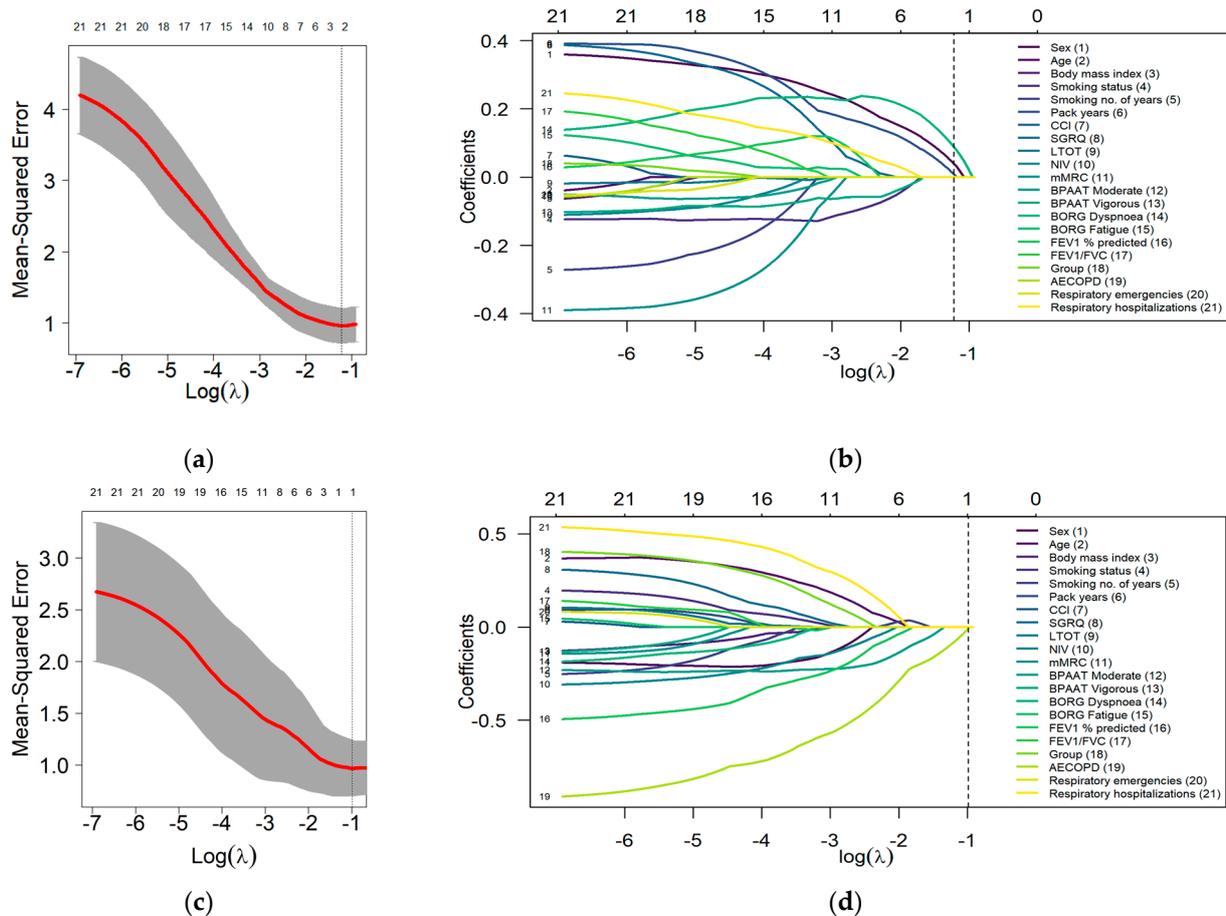


Figure A5. Cont.

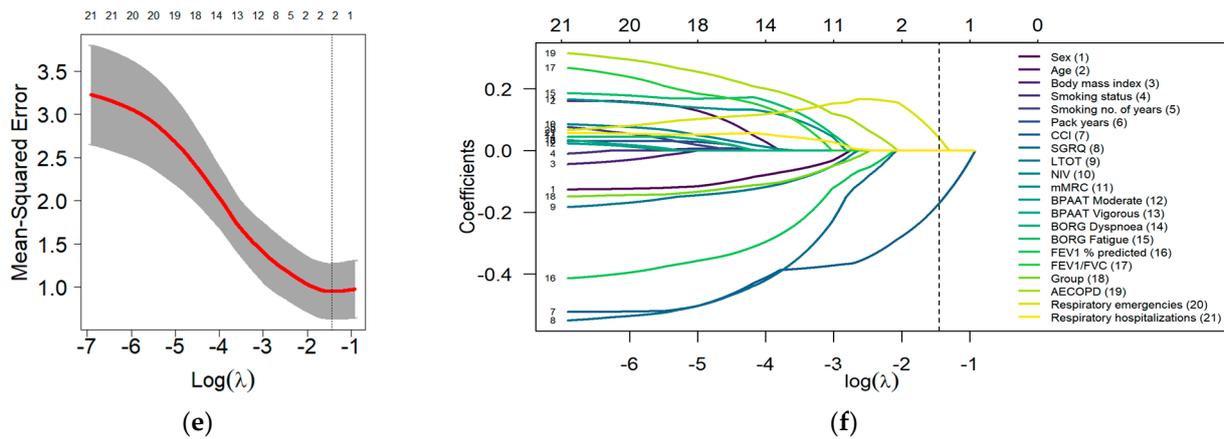


Figure A5. LASSO’s distribution of the 5-folds cross-validation mean squared error for the difference in the (a) handgrip muscle strength, (c) one-minute sit-to-stand test and (e) COPD assessment test values. Coefficients as a function of the natural logarithm of the penalty parameter λ for the difference in the (b) handgrip muscle strength, (d) one-minute sit-to-stand test and (f) COPD assessment test values. The minimum value of $\log(\lambda)$ is indicated by a vertical dotted line. Abbreviations: AECOPD, acute exacerbation of COPD; BPAAT, brief physical activity assessment tool; CCI, Charlson comorbidity index; COPD, chronic obstructive pulmonary disease; FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity; LTOT, long-term oxygen therapy; mMRC, modified medical council dyspnoea scale; NIV, non-invasive ventilation; no., number; SGRQ, St. George’s respiratory questionnaire.

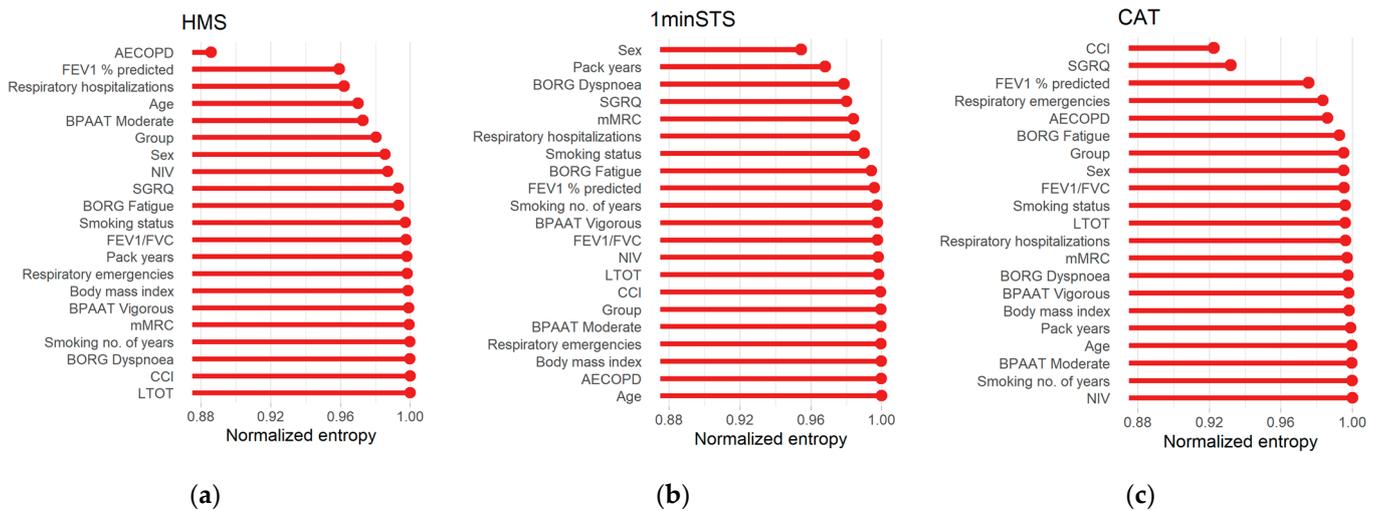


Figure A6. Feature importance given by the normalized entropy algorithm for: (a) the difference in the handgrip muscle strength (HMS); (b) the one-minute sit-to-stand test (1minSTS); (c) COPD assessment test (CAT) ($n = 38; 39; 42$). Abbreviations: AECOPD, acute exacerbation of COPD; BPAAT, brief physical activity assessment tool; BMI, body mass index; CCI, Charlson comorbidity index; COPD, chronic obstructive pulmonary disease; FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity; LTOT, long-term oxygen therapy; mMRC, modified medical council dyspnoea scale; NIV, non-invasive ventilation; no., number; SGRQ, St. George’s respiratory questionnaire.

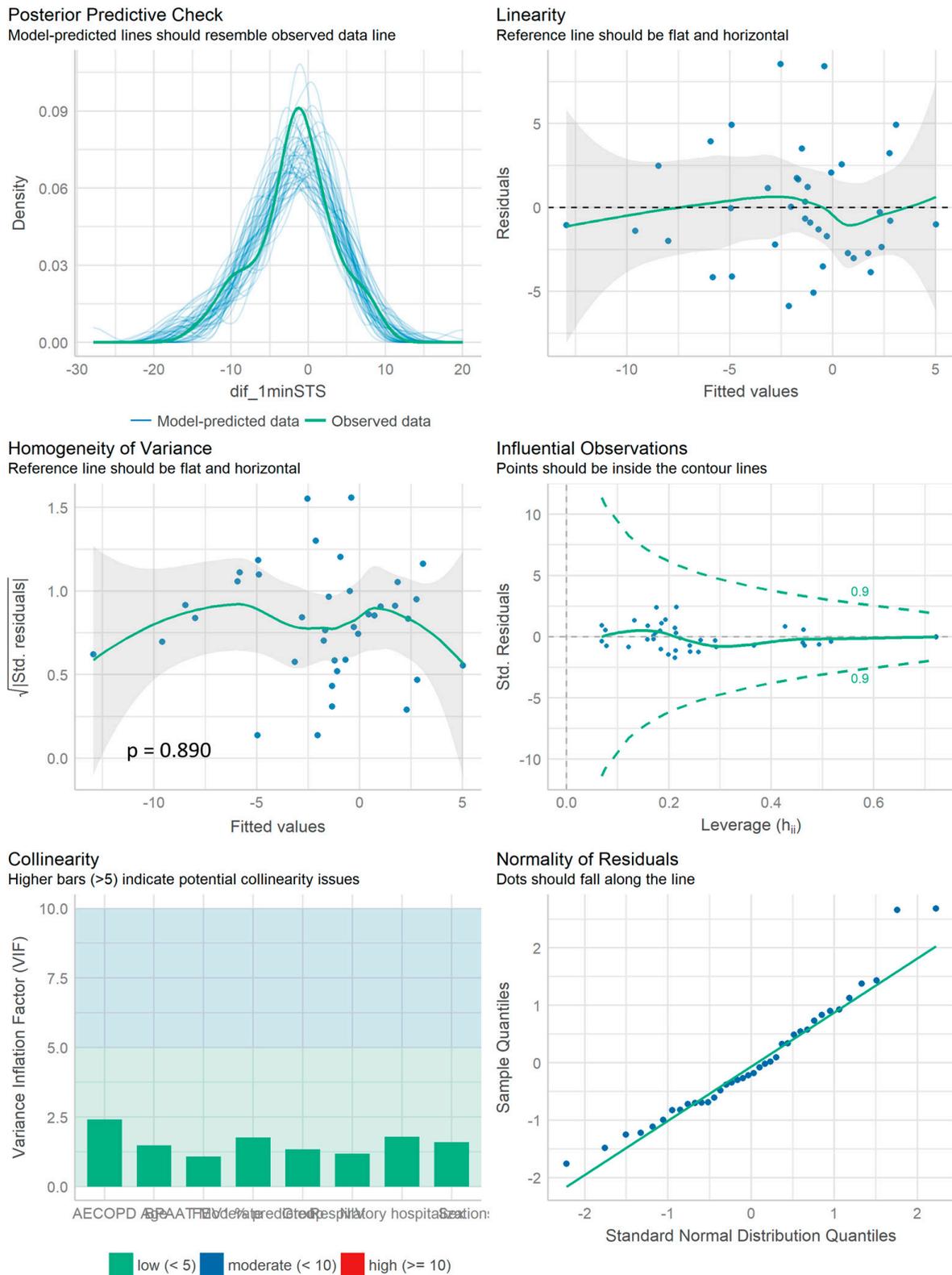


Figure A7. Residual analysis for the linear model using as dependent variable the difference in the handgrip muscle strength (dHMS) and the 8 most important features in people with chronic obstructive pulmonary disease ($n = 38$). Abbreviations: p , p -value for the Breusch-Pagan test.

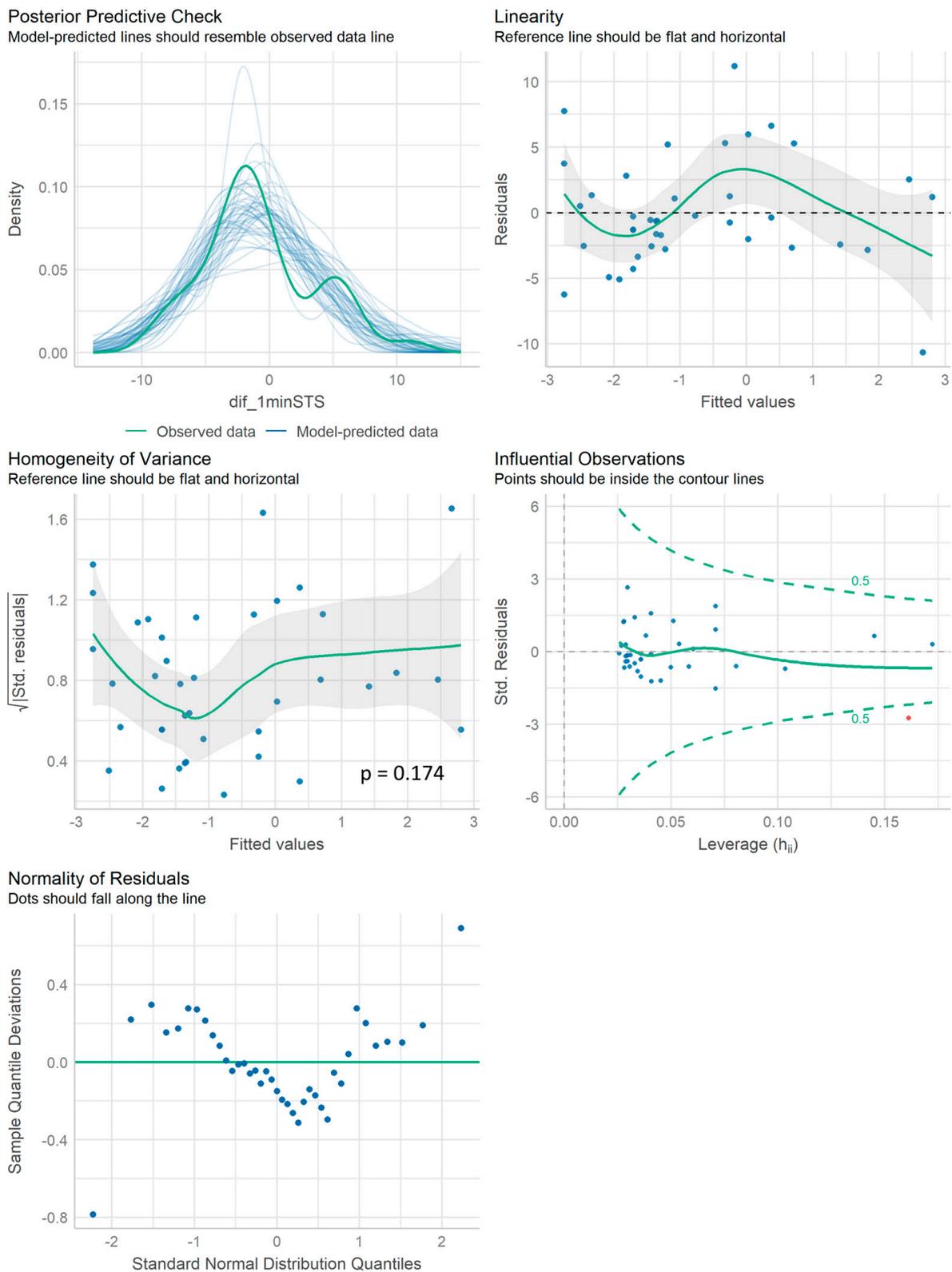


Figure A8. Residual analysis for the linear model using as dependent variable the difference in the number of repetitions in the one-minute sit-to-stand test (d1minSTS) and the most important feature in people with chronic obstructive pulmonary disease ($n = 39$). Abbreviations: p , p -value for the Breusch-Pagan test.

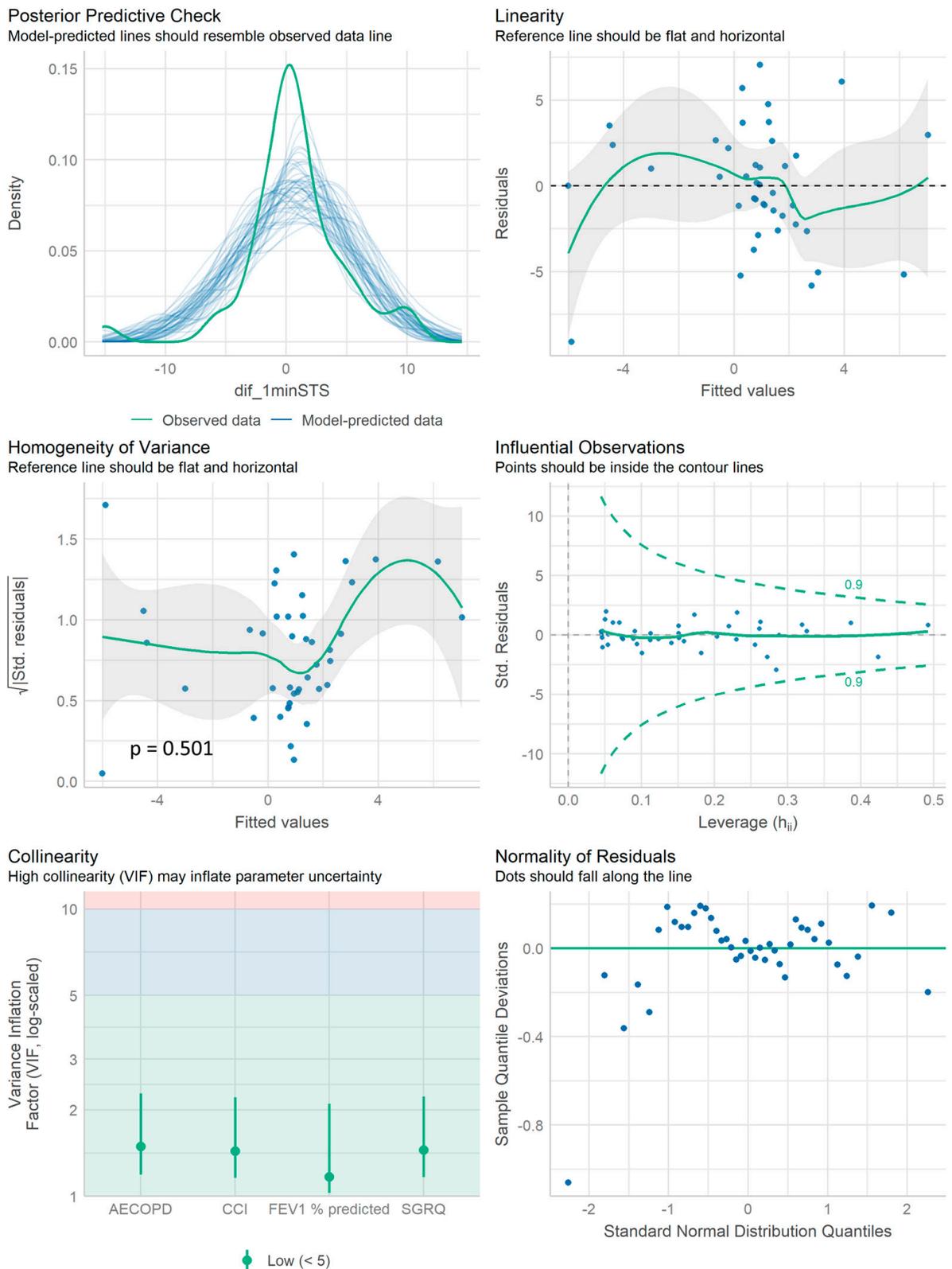


Figure A9. Residual analysis for the linear model using as dependent variable the difference in the COPD assessment test score (dCAT) and the 4 most important features in people with chronic obstructive pulmonary disease ($n = 42$). Abbreviations: p , p -value for the Breusch-Pagan test.

Table A1. Results from the hyperparameters tuning for the extreme gradient boosting algorithm for the difference in the handgrip muscle strength, the one-minute sit-to-stand test and the COPD assessment test values in people with chronic obstructive pulmonary disease (COPD) ($n = 38; 39; 42$). Note: Only the 10 lowest minimum RMSE values in the test set are presented for each outcome measure.

	eta	Maximum Tree Depth	Minimum Child Weight	Subsample Ratio	Train Set		Test Set	
					Iteration Number	Minimum RMSE	Iteration Number	Minimum RMSE
HMS	0.025	5	1	0.4	750	0.042563	105	1.017910
	0.025	8	1	0.4	750	0.041492	211	1.020147
	0.025	10	1	0.4	750	0.041492	211	1.020147
	0.025	11	1	0.4	750	0.041492	211	1.020147
	0.025	12	1	0.4	750	0.041492	211	1.020147
	0.025	14	1	0.4	750	0.041492	211	1.020147
	0.025	17	1	0.4	750	0.041492	211	1.020147
	0.015	5	1	0.4	750	0.126912	219	1.025573
	0.015	8	1	0.4	750	0.126209	219	1.025935
	0.015	10	1	0.4	750	0.126221	219	1.025935
1minSTS	0.020	5	3	0.6	750	0.067614	61	1.004571
	0.020	8	3	0.6	750	0.067625	61	1.004571
	0.020	10	3	0.6	750	0.067625	61	1.004571
	0.020	11	3	0.6	750	0.067625	61	1.004571
	0.020	12	3	0.6	750	0.067625	61	1.004571
	0.020	14	3	0.6	750	0.067625	61	1.004571
	0.020	17	3	0.6	750	0.067625	61	1.004571
	0.010	8	2	0.6	750	0.131015	135	1.011578
	0.010	10	2	0.6	750	0.131015	135	1.011578
	0.010	11	2	0.6	750	0.131015	135	1.011578
CAT	0.020	5	3	0.6	750	0.124623	52	1.015119
	0.020	8	3	0.6	750	0.124178	52	1.015119
	0.020	10	3	0.6	750	0.124178	52	1.015119
	0.020	11	3	0.6	750	0.124178	52	1.015119
	0.020	12	3	0.6	750	0.124178	52	1.015119
	0.020	14	3	0.6	750	0.124178	52	1.015119
	0.020	17	3	0.6	750	0.124178	52	1.015119
	0.025	5	3	0.6	750	0.085357	29	1.015720
	0.025	8	3	0.6	750	0.085024	29	1.015720
	0.025	10	3	0.6	750	0.085024	29	1.015720

Abbreviations: 1minSTS, one-minute sit-to-stand test; CAT, COPD assessment test; eta, learning rate; HMS, handgrip muscle strength; RMSE, root mean squared error.

References

1. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006.
2. Jobson, J.D. (Ed.) *Multiple Linear Regression BT—Applied Multivariate Data Analysis: Regression and Experimental Design*; Springer: New York, NY, USA, 1991; pp. 219–398, ISBN 978-1-4612-0955-3.
3. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer Science & Business Media: New York, NY, USA, 2009; ISBN 0387848584.
4. Abu-Mostafa, Y.S.; Magdon-Ismael, M.; Lin, H.-T. *Learning from Data*; AMLBook: New York, NY, USA, 2012; Volume 4.
5. Gareth, J.; Hastie, T.; Tibshirani, R.; Witten, D. *An Introduction to Statistical Learning: With Applications in R*; Springer Science + Business Media, LLC: New York, NY, USA, 2013.
6. George, E.I. The Variable Selection Problem. *J. Am. Stat. Assoc.* **2000**, *95*, 1304–1308. [[CrossRef](#)]
7. Guyon, I.; Elisseeff, A. An Introduction to Variable and Feature Selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.
8. Liu, S.; Yao, J.; Zhou, C.; Motani, M. SURI: Feature Selection Based on Unique Relevant Information for Health Data. In *Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Madrid, Spain, 3–6 December 2018; pp. 687–692.
9. Fan, J.; Li, R. Variable Selection via Nonconcave Penalized Likelihood and Its Oracle Properties. *J. Am. Stat. Assoc.* **2001**, *96*, 1348–1360. [[CrossRef](#)]
10. Lin, D.; Foster, D.P.; Ungar, L.H. VIF Regression: A Fast Regression Algorithm for Large Data. *J. Am. Stat. Assoc.* **2011**, *106*, 232–247. [[CrossRef](#)]

11. Ambroise, C.; McLachlan, G.J. Selection Bias in Gene Extraction on the Basis of Microarray Gene-Expression Data. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6562–6566. [[CrossRef](#)] [[PubMed](#)]
12. Weisberg, S. *Applied Linear Regression*, 4th ed.; Wiley: New Jersey, NJ, USA, 2013.
13. Whittingham, M.J.; Stephens, P.A.; Bradbury, R.B.; Freckleton, R.P. Why Do We Still Use Stepwise Modelling in Ecology and Behaviour? *J. Anim. Ecol.* **2006**, *75*, 1182–1189. [[CrossRef](#)] [[PubMed](#)]
14. Smith, G. Step Away from Stepwise. *J. Big Data* **2018**, *5*, 32. [[CrossRef](#)]
15. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
16. Kursa, M.; Jankowski, A.; Rudnicki, W. Boruta—A System for Feature Selection. *Fundam. Inf.* **2010**, *101*, 271–285. [[CrossRef](#)]
17. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 8 March 2016; Volume 13–17, pp. 785–794.
18. Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B (Methodol.)* **1996**, *58*, 267–288. [[CrossRef](#)]
19. Jaynes, E.T. Information Theory and Statistical Mechanics. *Phys. Rev.* **1957**, *106*, 620–630. [[CrossRef](#)]
20. Golan, A. *Foundations of Info-Metrics*; Oxford University Press: Oxford, UK, 2017; Volume 1, ISBN 9780199349524.
21. Chen, M.; Dunn, J.M.; Golan, A.; Ullah, A. (Eds.) *Advances in Info-Metrics*; Oxford University Press: Oxford, UK, 2020; ISBN 9780190636685.
22. Mittelhammer, R.; Cardell, N.; Marsh, T. The Data-Constrained Generalized Maximum Entropy Estimator of the GLM: Asymptotic Theory and Inference. *Entropy* **2013**, *15*, 1756–1775. [[CrossRef](#)]
23. Golan, A.; Judge, G.G.; Miller, D. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*; Wiley: Chichester, UK; New York, NY, USA, 1996; ISBN 0471953113 9780471953111.
24. Satheeshkumar, P.S.; El-Dallal, M.; Mohan, M.P. Feature Selection and Predicting Chemotherapy-Induced Ulcerative Mucositis Using Machine Learning Methods. *Int. J. Med. Inform.* **2021**, *154*, 104563. [[CrossRef](#)]
25. Hall, M.-H.; Holton, K.M.; Öngür, D.; Montrose, D.; Keshavan, M.S. Longitudinal Trajectory of Early Functional Recovery in Patients with First Episode Psychosis. *Schizophr. Res.* **2019**, *209*, 234–244. [[CrossRef](#)]
26. Kiley, J.P.; Sri Ram, J.; Croxton, T.L.; Weinmann, G.G. Challenges Associated with Estimating Minimal Clinically Important Differences in COPD—The NHLBI Perspective. *COPD J. Chronic Obst. Pulm. Dis.* **2005**, *2*, 43–46. [[CrossRef](#)] [[PubMed](#)]
27. Global Initiative for Chronic Obstructive Lung Disease GOLD Report 2023. In *Global Initiative for Chronic Obstructive Lung Disease*; Global Initiative for Chronic Obstructive Lung Disease, Inc.: Madison, WI, USA, 2023.
28. Levine, S.M.; Marciniuk, D.D. Global Impact of Respiratory Disease: What Can We Do, Together, to Make a Difference? *Chest* **2022**, *161*, 1153–1154. [[CrossRef](#)] [[PubMed](#)]
29. Momtazmanesh, S.; Moghaddam, S.S.; Ghamari, S.-H.; Rad, E.M.; Rezaei, N.; Shobeiri, P.; Aali, A.; Abbasi-Kangevari, M.; Abbasi-Kangevari, Z.; Abdelmasseh, M.; et al. Global Burden of Chronic Respiratory Diseases and Risk Factors, 1990–2013; 2019: An Update from the Global Burden of Disease Study 2019. *eClinicalMedicine* **2023**, *59*, 101936. [[CrossRef](#)] [[PubMed](#)]
30. Varmaghani, M.; Dehghani, M.; Heidari, E.; Sharifi, F.; Moghaddam, S.S.; Farzadfar, F. Global Prevalence of Chronic Obstructive Pulmonary Disease: Systematic Review and Meta-Analysis. *East. Mediterr. Health J.* **2019**, *25*, 47–57. [[CrossRef](#)] [[PubMed](#)]
31. Jarad, N. Chronic Obstructive Pulmonary Disease (COPD) and Old Age? *Chronic Respir. Dis.* **2011**, *8*, 143–151. [[CrossRef](#)]
32. Rennard, S.I.; Drummond, M.B. Early Chronic Obstructive Pulmonary Disease: Definition, Assessment, and Prevention. *Lancet* **2015**, *385*, 1778–1788. [[CrossRef](#)]
33. Sun, Y.; Milne, S.; Jaw, J.E.; Yang, C.X.; Xu, F.; Li, X.; Obeidat, M.; Sin, D.D. BMI Is Associated with FEV1 Decline in Chronic Obstructive Pulmonary Disease: A Meta-Analysis of Clinical Trials. *Respir. Res.* **2019**, *20*, 236. [[CrossRef](#)] [[PubMed](#)]
34. Cao, C.; Wang, R.; Wang, J.; Bunjhoo, H.; Xu, Y.; Xiong, W. Body Mass Index and Mortality in Chronic Obstructive Pulmonary Disease: A Meta-Analysis. *PLoS ONE* **2012**, *7*, e43892. [[CrossRef](#)] [[PubMed](#)]
35. Acharya, V.K.; Sharma, D.K.; Kamath, S.K.; Shreenivasa, A.; Unnikrishnan, B.; Holla, R.; Gautham, M.; Rathi, P.; Mendonca, J. Impact of COVID-19 Pandemic on the Exacerbation Rates in COPD Patients in Southern India—A Potential Role for Community Mitigations Measures. *Int. J. Chronic Obstruct. Pulm. Dis.* **2023**, *18*, 1909–1917. [[CrossRef](#)] [[PubMed](#)]
36. Alsallakh, M.A.; Sivakumaran, S.; Kennedy, S.; Vasileiou, E.; Lyons, R.A.; Robertson, C.; Sheikh, A.; Davies, G.A.; Simpson, C.R.; McMenamin, J.; et al. Impact of COVID-19 Lockdown on the Incidence and Mortality of Acute Exacerbations of Chronic Obstructive Pulmonary Disease: National Interrupted Time Series Analyses for Scotland and Wales. *BMC Med.* **2021**, *19*, 124. [[CrossRef](#)] [[PubMed](#)]
37. Nishioki, T.; Sato, T.; Okajima, A.; Motomura, H.; Takeshige, T.; Watanabe, J.; Yae, T.; Koyama, R.; Kido, K.; Takahashi, K. Impact of the COVID-19 Pandemic on COPD Exacerbations in Japanese Patients: A Retrospective Study. *Sci. Rep.* **2024**, *14*, 2792. [[CrossRef](#)] [[PubMed](#)]
38. González, J.; Moncusí-Moix, A.; Benítez, I.D.; Santistevé, S.; Monge, A.; Fontiveros, M.A.; Carmona, P.; Torres, G.; Barbé, F.; de Batlle, J. Clinical Consequences of COVID-19 Lockdown in Patients With COPD: Results of a Pre-Post Study in Spain. *Chest* **2021**, *160*, 135–138. [[CrossRef](#)] [[PubMed](#)]
39. Bakaloudi, D.R.; Barazzoni, R.; Bischoff, S.C.; Breda, J.; Wickramasinghe, K.; Chourdakis, M. Impact of the First COVID-19 Lockdown on Body Weight: A Combined Systematic Review and a Meta-Analysis. *Clin. Nutr.* **2022**, *41*, 3046–3054. [[CrossRef](#)] [[PubMed](#)]
40. Siu, H.; Polkinghorne, K.; Finlay, P.; Yong, T.; Bardin, P.G.; King, P.T. Effect of COVID-19 Lockdown on Body Weight in Chronic Obstructive Pulmonary Disease. *Intern. Med. J.* **2023**, *53*, 615–618. [[CrossRef](#)] [[PubMed](#)]

41. Charlson, M.; Szatrowski, T.P.; Peterson, J.; Gold, J. Validation of a Combined Comorbidity Index. *J. Clin. Epidemiol.* **1994**, *47*, 1245–1251. [[CrossRef](#)]
42. Graham, B.L.; Steenbruggen, I.; Barjaktarevic, I.Z.; Cooper, B.G.; Hall, G.L.; Hallstrand, T.S.; Kaminsky, D.A.; McCarthy, K.; McCormack, M.C.; Miller, M.R.; et al. Standardization of Spirometry 2019 Update an Official American Thoracic Society and European Respiratory Society Technical Statement. *Am. J. Respir. Crit. Care Med.* **2019**, *200*, E70–E88. [[CrossRef](#)]
43. Crisafulli, E.; Clini, E.M. Measures of Dyspnea in Pulmonary Rehabilitation. *Multidiscip. Respir. Med.* **2010**, *5*, 202. [[CrossRef](#)] [[PubMed](#)]
44. Bestall, J.C.; Paul, E.A.; Garrod, R.; Garnham, R.; Jones, P.W.; Wedzicha, J.A. Usefulness of the Medical Research Council (MRC) Dyspnoea Scale as a Measure of Disability in Patients with Chronic Obstructive Pulmonary Disease. *Thorax* **1999**, *54*, 581–586. [[CrossRef](#)] [[PubMed](#)]
45. Mahler, D.A.; Rosiello, R.A.; Harver, A.; Lentine, T.; McGovern, J.F.; Daubenspeck, J.A. Comparison of Clinical Dyspnea Ratings and Psychophysical Measurements of Respiratory Sensation in Obstructive Airway Disease. *Am. Rev. Respir. Dis.* **1987**, *135*, 1229–1233. [[CrossRef](#)] [[PubMed](#)]
46. Wilson, R.C.; Jones, P.W. A Comparison of the Visual Analogue Scale and Modified Borg Scale for the Measurement of Dyspnoea during Exercise. *Clin. Sci.* **1989**, *76*, 277–282. [[CrossRef](#)] [[PubMed](#)]
47. Borg, G.A. Psychophysical Bases of Perceived Exertion. *Med. Sci. Sports Exerc.* **1982**, *14*, 377–381. [[CrossRef](#)] [[PubMed](#)]
48. Marshall, A.L.; Smith, B.J.; Bauman, A.E.; Kaur, S. Reliability and Validity of a Brief Physical Activity Assessment for Use by Family Doctors. *Br. J. Sports Med.* **2005**, *39*, 294–297. [[CrossRef](#)] [[PubMed](#)]
49. Jones, P.W.; Quirk, F.H.; Baveystock, C.M. The St George’s Respiratory Questionnaire. *Respir. Med.* **1991**, *85* (Suppl. SB), 25–27. [[CrossRef](#)]
50. Clegg, A.; Young, J.; Iliffe, S.; Rikkert, M.O.; Rockwood, K. Frailty in Elderly People. *Lancet* **2013**, *381*, 752–762. [[CrossRef](#)] [[PubMed](#)]
51. Vaidya, T.; Chambellan, A.; de Bisschop, C. Sit-to-Stand Tests for COPD: A Literature Review. *Respir. Med.* **2017**, *128*, 70–77. [[CrossRef](#)]
52. Ozalevli, S.; Ozden, A.; Itil, O.; Akkoçlu, A. Comparison of the Sit-to-Stand Test with 6 Min Walk Test in Patients with Chronic Obstructive Pulmonary Disease. *Respir. Med.* **2007**, *101*, 286–293. [[CrossRef](#)]
53. Bohannon, R.W. Minimal Clinically Important Difference for Grip Strength: A Systematic Review. *J. Phys. Ther. Sci.* **2019**, *31*, 75–78. [[CrossRef](#)]
54. Vaidya, T.; de Bisschop, C.; Beaumont, M.; Ouksel, H.; Jean, V.; Dessables, F.; Chambellan, A. Is the 1-Minute Sit-to-Stand Test a Good Tool for the Evaluation of the Impact of Pulmonary Rehabilitation? Determination of the Minimal Important Difference in COPD. *Int. J. Chronic Obstruct. Pulmon. Dis.* **2016**, *11*, 2609–2616. [[CrossRef](#)]
55. George, F. *Diagnóstico e Tratamento Da Doença Pulmonar Obstrutiva Crónica*; 028/2011; Direção Geral da Saúde: Lisbon, Portugal, 2013.
56. Jones, P.W.; Harding, G.; Berry, P.; Wiklund, I.; Chen, W.-H.; Kline Leidy, N. Development and First Validation of the COPD Assessment Test. *Eur. Respir. J.* **2009**, *34*, 648. [[CrossRef](#)] [[PubMed](#)]
57. Kon, S.S.C.; Canavan, J.L.; Jones, S.E.; Nolan, C.M.; Clark, A.L.; Dickson, M.J.; Haselden, B.M.; Polkey, M.I.; Man, W.D.-C. Minimum Clinically Important Difference for the COPD Assessment Test: A Prospective Analysis. *Lancet Respir. Med.* **2014**, *2*, 195–203. [[CrossRef](#)]
58. Akaike, H. Maximum Likelihood Identification of Gaussian Autoregressive Moving Average Models. *Biometrika* **1973**, *60*, 255–265. [[CrossRef](#)]
59. Schwarz, G. Estimating the Dimension of a Model. *Ann. Stat.* **1978**, *6*, 461–464. [[CrossRef](#)]
60. Tibshirani, R. *Bias, Variance, and Prediction Error for Classification Rules*; University of Toronto: Toronto, ON, Canada, 1996.
61. Breiman, L. Bagging Predictors. *Mach Learn.* **1996**, *24*, 123–140. [[CrossRef](#)]
62. Chen, T.; He, T.; Benesty, M.; Khotilovich, V.; Tang, Y.; Cho, H.; Chen, K.; Mitchell, R.; Cano, I.; Zhou, T.; et al. *Xgboost: Extreme Gradient Boosting*. 2021. R Package Version 1.7.7.1. 2024. Available online: <https://CRAN.R-project.org/package=xgboost> (accessed on 15 February 2024).
63. Zuur, A.; Ieno, E.; Walker, N.; Saveliev, A.; Smith, G. *Mixed Effects Models and Extensions in Ecology With R*; Springer: New York, NY, USA, 2009.
64. Macedo, P. Freedman’s Paradox: A Solution Based on Normalized Entropy. In *Theory and Applications of Time Series Analysis, Proceedings of the ITISE 2019, Granada, Spain, 20–27 September 2019*; Valenzuela, O., Rojas, F., Herrera, L.J., Pomares, H., Rojas, I., Eds.; Springer: New York, NY, USA, 2020; pp. 239–252.
65. Macedo, P.; Costa, M.C.; Cruz, J.P. Normalized Entropy: A Comparison with Traditional Techniques in Variable Selection. *AIP Conf. Proc.* **2022**, *2425*, 190002.
66. Hoerl, A.E.; Kennard, R.W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970**, *12*, 55–67. [[CrossRef](#)]
67. KENDALL, M.G. A NEW MEASURE OF RANK CORRELATION. *Biometrika* **1938**, *30*, 81–93. [[CrossRef](#)]
68. KENDALL, M.G. THE TREATMENT OF TIES IN RANKING PROBLEMS. *Biometrika* **1945**, *33*, 239–251. [[CrossRef](#)] [[PubMed](#)]
69. Burnham, K.P.; Anderson, D.R. *Model Selection and Multimodel Inference*, 2nd ed.; Springer: New York, NY, USA, 2002; ISBN 978-0-387-95364-9.

70. Lüdecke, D. Ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. *J. Open Source Softw.* **2018**, *3*, 772. [CrossRef]
71. Wiley, J.F. *JWileymisc: Miscellaneous Utilities and Functions*. 2022. R Package Version 1.4.1. 2023. Available online: <https://CRAN.R-project.org/package=JWileymisc> (accessed on 15 February 2024).
72. Ishwaran, H.; Kogalur, U.B. *Fast Unified Random Forests for Survival, Regression, and Classification (RF-SRC)*. 2021. R Package Version 3.2.3. 2023. Available online: <https://CRAN.R-project.org/package=randomForestSRC> (accessed on 15 February 2024).
73. Liaw, A.; Wiener, M. Classification and Regression by RandomForest. *R. News* **2002**, *2*, 18–22.
74. Kursa, M.B.; Rudnicki, W.R. Feature Selection with the Boruta Package. *J. Stat. Softw.* **2010**, *36*, 1–13. [CrossRef]
75. Friedman, J.H.; Hastie, T.; Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **2010**, *33*, 1–22. [CrossRef] [PubMed]
76. Venables, W.N.; Ripley, B.D. *Modern Applied Statistics with S*; Springer: New York, NY, USA, 2002; ISBN 0387954570, 9780387954578, 9781441930088, 1441930086.
77. Lüdecke, D.; Ben-Shachar, M.S.; Patil, I.; Waggoner, P.; Makowski, D. Performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *J. Open Source Softw.* **2021**, *6*, 3139. [CrossRef]
78. Lüdecke, D. *SjPlot: Data Visualization for Statistics in Social Science*. 2021. R Package Version 2.8.15. 2023. Available online: <https://CRAN.R-project.org/package=sjPlot> (accessed on 15 February 2024).
79. RStudio Team. *RStudio: Integrated Development Environment for R*. 2023. Version 2023.12.1+402. 2023. Available online: <https://posit.co/> (accessed on 15 February 2024).
80. R Core Team. *R: A Language and Environment for Statistical Computing*. 2023. Version 4.3.3. 2023. Available online: <https://www.r-project.org/> (accessed on 15 February 2024).
81. Hasan, N.; Bao, Y. Comparing Different Feature Selection Algorithms for Cardiovascular Disease Prediction. *Health Technol.* **2021**, *11*, 49–62. [CrossRef]
82. Freedman, D.A. A Note on Screening Regression Equations. *Am. Stat.* **1983**, *37*, 152–155. [CrossRef]
83. He, H.; Jin, H.; Chen, J. Automatic Feature Selection for Classification of Health Data. In Proceedings of the AI 2005: Advances in Artificial Intelligence, AI 2005, Sydney, Australia, 5–9 December 2005; Zhang, S., Jarvis, R., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; pp. 910–913.
84. Afreixo, V.; Cabral, J.; Macedo, P. Comparison of Feature Selection Methods in Regression Modeling: A Simulation Study. In Proceedings of the Computational Science and Its Applications—ICCSA 2023 Workshops, ICCSA 2023, Athens, Greece, 3–6 July 2023; Gervasi, O., Murgante, B., Rocha, A.M.A.C., Garau, C., Scorza, F., Karaca, Y., Torre, C.M., Eds.; Springer Nature: Cham, Switzerland, 2023; pp. 150–159.
85. Rassouli, F.; Baty, F.; Stolz, D.; Albrich, W.; Tamm, M.; Widmer, S.; Brutsche, M. Longitudinal Change of COPD Assessment Test (CAT) in a Telehealthcare Cohort Is Associated with Exacerbation Risk. *Int. J. Chronic Obstruct. Pulmon. Dis.* **2017**, *12*, 3103–3109. [CrossRef]
86. Feng, J.; Liang, J.; Qiang, Z.; Li, X.; Chen, Q.; Liu, G.; Hong, J.; Hao, Z.; Wei, H. Effective Techniques for Intelligent Cardiocography Interpretation Using XGB-RF Feature Selection and Stacking Fusion. In Proceedings of the 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Houston, TX, USA, 9–12 December 2021; pp. 2667–2673.
87. Xu, Z.; Wang, Z. A Risk Prediction Model for Type 2 Diabetes Based on Weighted Feature Selection of Random Forest and XGBoost Ensemble Classifier. In Proceedings of the 2019 Eleventh International Conference on Advanced Computational Intelligence (ICACI), Guilin, China, 7–9 June 2019; pp. 278–283.
88. Wiegand, R.E. Performance of Using Multiple Stepwise Algorithms for Variable Selection. *Stat. Med.* **2010**, *29*, 1647–1659. [CrossRef]
89. Kumar, S.S.; Shaikh, T. Empirical Evaluation of the Performance of Feature Selection Approaches on Random Forest. In Proceedings of the 2017 International Conference on Computer and Applications (ICCA), New York, NY, USA, 21 April 2017; pp. 227–231.
90. Sanchez-Pinto, L.N.; Venable, L.R.; Fahrenbach, J.; Churpek, M.M. Comparison of Variable Selection Methods for Clinical Predictive Modeling. *Int. J. Med. Inform.* **2018**, *116*, 10–17. [CrossRef] [PubMed]

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.