

Modelo de Inteligencia Artificial – Machine Learning explicable para la predicción de sepsis.

Miguel Angel Duarte Millán

Anexo 1.

1. Rangos definidos para las variables.
2. Variables seleccionadas para el entrenamiento.
 - 3.1 Rendimiento de modelos RF balanceados, filtrados por Sensibilidad > 0.80 .
 - 3.2 Rendimiento de modelos XGB balanceados, filtrados por Sensibilidad > 0.80 .
 - 4.1 *Permutation importance* de modelo RF elegido.
 - 4.2 *Permutation importance* de modelo XGB elegido.

1. Rangos definidos para las variables.

Parámetro	Mínimo	Máximo
Actividad de protrombina (%)	10	125
Alanina aminotransferasa (GPT) (U/L)	1	15000
Albúmina (g/dL)	0.50	7
Aspartato aminotransferasa (AST) (U/L)	1	15000
Bicarbonato arterial (mEq/L)	5	70
Bicarbonato venoso (mEq/L)	5	70
Bilirrubina total (mg/dL)	0.10	30
Creatinina (mg/dL)	0.10	20
Fibrinógeno calculado (mg/dL)	30	5000
Flujo de oxígeno (L/min)	0	50
Frecuencia cardíaca (lpm)	30	200
Frecuencia respiratoria (rpm)	5	50
Gammaglutamiltransferasa (U/L)	1	15000
Glucosa (mg/dL)	30	1000
Hemoglobina (g/dL)	3	20
Lactato arterial (mmol/L)	0.10	20
Lactato venoso (mmol/L)	0.10	20
Leucocitos (por μL)	0.01	40
Linfocitos ($10^3/\mu\text{L}$)	0.01	20
Neutrófilos ($10^3/\mu\text{L}$)	0.01	30
pH arterial (sin unidad)	6.80	7.80
pH venoso (sin unidad)	6.80	7.80
Plaquetas ($10^3/\mu\text{L}$)	5	1000
Porcentaje de neutrófilos (%)	0.1	99
Potasio (mEq/L)	1.50	10
Presión parcial de CO_2 arterial (mmHg)	10	150
Presión parcial de CO_2 venosa (mmHg)	10	150
Presión parcial de O_2 arterial (mmHg)	10	150
Presión parcial de O_2 venosa (mmHg)	10	150
Procalcitonina (ng/mL)	0.001	100
Proteína C reactiva (mg/dL)	0.001	100
Puntuación de Glasgow total (escala 3–15)	3	15
Saturación de O_2 (%)	70	100
Saturación de O_2 arterial (%)	20	100
Saturación de O_2 venosa (%)	20	100
Sodio (mEq/L)	100	190
Temperatura ($^{\circ}\text{C}$)	33	42
Tensión arterial diastólica (mmHg)	30	150
Tensión arterial sistólica (mmHg)	50	250
Tiempo de tromboplastina parcial activado (s)	15	65
Urea (mg/dL)	5	500

2. Variables seleccionadas para el entrenamiento:

Demográficas	Comorbilidades (CIE-10)	Constantes
Sexo (H,M)	Cardiopatía isquémica (I20–I25)	Temperatura (°C)
Edad (Años)	Insuficiencia cardíaca (I50)	Tensión arterial sistólica (mm Hg)
	Arteriopatía periférica (I70.2)	Tensión arterial diastólica (mm Hg)
	EPOC (J44)	Frecuencia cardíaca (lpm)
	Asma (J45)	Saturación periférica O ₂ . (%)
	Demencia (F02)	Flujo O ₂ (L/min)
	Alzheimer (G20)	
	Enfermedad oncológica (C00–C80)	
	Enfermedad hematológica (C81–C96)	
	Enfermedad autoinmune (M30–M36, I77.82)	
	Hipertensión (I10)	
	Diabetes (E10–E14)	
Laboratorio		
Actividad de protrombina (%)	Alanina aminotransferasa (GPT) (U/L)	Albúmina (g/dL)
Aspartato aminotransferasa (AST) (U/L)	Bicarbonato arterial (mEq/L)	Bicarbonato venoso (mEq/L)
Bilirrubina total (mg/dL)	Creatinina (mg/dL)	Fibrinógeno calculado (mg/dL)
Gamma-glutamilttransferasa (U/L)	Glucosa (mg/dL)	Hemoglobina (g/dL)
Lactato arterial (mmol/L)	Lactato venoso (mmol/L)	Leucocitos (por μ L)
Linfocitos ($10^3/\mu$ L)	Neutrófilos ($10^3/\mu$ L)	pH arterial (sin unidad)
pH venoso (sin unidad)	Plaquetas ($10^3/\mu$ L)	Porcentaje de neutrófilos (%)
Potasio (mEq/L)	Presión parcial de CO ₂ arterial (mmHg)	Presión parcial de CO ₂ venosa (mmHg)
Presión parcial de O ₂ arterial (mmHg)	Presión parcial de O ₂ venosa (mmHg)	Procalcitonina (ng/mL)
Proteína C reactiva (mg/dL)	Saturación de O ₂ arterial (%)	Saturación de O ₂ venosa (%)
Sodio (mEq/L)	Tiempo de tromboplastina parcial activado (s)	Urea (mg/dL)

3.1 Rendimiento de modelos RF balanceados, filtrados por Sensibilidad > 0.80:

modelo	threshold	roc_auc	auprc	accuracy	precision	recall	specificity	f1
Med_US RF n=200 d=8	0,4	0,83	0,34	0,60	0,19	0,90	0,57	0,31
Med_US RF n=300 d=8	0,4	0,83	0,34	0,60	0,19	0,90	0,57	0,31
Med_US RF n=200 d=10	0,4	0,84	0,35	0,62	0,19	0,89	0,59	0,32
KNN_US RF n=200 d=8	0,4	0,82	0,32	0,59	0,18	0,89	0,56	0,30
Med_US RF n=300 d=12	0,4	0,84	0,35	0,64	0,20	0,89	0,61	0,33
Med_US RF n=200 d=12	0,4	0,84	0,35	0,64	0,20	0,89	0,61	0,33
KNN_US RF n=300 d=12	0,4	0,83	0,32	0,62	0,19	0,89	0,59	0,32
KNN_US RF n=200 d=10	0,4	0,82	0,32	0,61	0,19	0,89	0,58	0,31
KNN_US RF n=300 d=8	0,4	0,82	0,32	0,59	0,18	0,89	0,55	0,30
KNN_US RF n=300 d=10	0,4	0,82	0,32	0,61	0,19	0,89	0,58	0,31
Med_US RF n=300 d=10	0,4	0,83	0,34	0,62	0,19	0,89	0,60	0,32
KNN_US RF n=200 d=12	0,4	0,83	0,32	0,62	0,19	0,89	0,59	0,32
Med_Over2 RF n=200 d=8	0,4	0,83	0,33	0,63	0,20	0,88	0,60	0,32
Med_Over2 RF n=300 d=8	0,4	0,83	0,33	0,63	0,20	0,88	0,60	0,32
Med_ROS RF n=200 d=8	0,4	0,83	0,33	0,63	0,20	0,88	0,60	0,32
KNN_Over2 RF n=300 d=8	0,4	0,82	0,31	0,62	0,19	0,88	0,59	0,31
KNN_Over2 RF n=200 d=8	0,4	0,82	0,31	0,62	0,19	0,88	0,59	0,31
KNN_ROS RF n=200 d=8	0,4	0,82	0,31	0,62	0,19	0,88	0,59	0,31
KNN_ROS RF n=300 d=8	0,4	0,82	0,31	0,61	0,19	0,87	0,59	0,31
Med_ROS RF n=300 d=8	0,4	0,83	0,33	0,63	0,19	0,87	0,60	0,32
Median RF n=300 d=8	0,4	0,83	0,33	0,65	0,20	0,87	0,62	0,33
Median RF n=200 d=8	0,4	0,83	0,32	0,65	0,20	0,87	0,63	0,33
KNN RF n=300 d=8	0,4	0,81	0,31	0,64	0,20	0,86	0,61	0,32
KNN RF n=200 d=8	0,4	0,81	0,31	0,64	0,20	0,86	0,61	0,32
Med_ROS RF n=200 d=10	0,4	0,83	0,33	0,68	0,22	0,86	0,66	0,35
Med_US RF n=300 d=12	0,45	0,84	0,35	0,69	0,22	0,86	0,67	0,35
Med_US RF n=200 d=8	0,45	0,83	0,34	0,67	0,21	0,86	0,65	0,34
Med_ROS RF n=300 d=10	0,4	0,83	0,33	0,68	0,22	0,85	0,66	0,34
Med_US RF n=200 d=12	0,45	0,84	0,35	0,69	0,22	0,85	0,67	0,35
Med_Over2 RF n=300 d=10	0,4	0,83	0,33	0,68	0,22	0,85	0,66	0,35
Med_US RF n=300 d=10	0,45	0,83	0,34	0,68	0,22	0,85	0,67	0,35
Med_US RF n=200 d=10	0,45	0,84	0,35	0,68	0,22	0,85	0,66	0,35
KNN_US RF n=200 d=12	0,45	0,83	0,32	0,68	0,21	0,85	0,66	0,34
KNN_US RF n=300 d=10	0,45	0,82	0,32	0,67	0,21	0,85	0,66	0,34
KNN_US RF n=200 d=8	0,45	0,82	0,32	0,66	0,21	0,85	0,64	0,33
Med_Over2 RF n=200 d=10	0,4	0,83	0,33	0,68	0,22	0,85	0,66	0,35
KNN_US RF n=300 d=12	0,45	0,83	0,32	0,68	0,21	0,85	0,66	0,34
Med_US RF n=300 d=8	0,45	0,83	0,34	0,67	0,21	0,85	0,65	0,34
KNN_US RF n=300 d=8	0,45	0,82	0,32	0,66	0,21	0,85	0,64	0,33

KNN_US RF n=200 d=10	0,45	0,82	0,32	0,67	0,21	0,84	0,65	0,34
KNN_ROS RF n=200 d=10	0,4	0,82	0,31	0,68	0,22	0,84	0,67	0,34
KNN_Over2 RF n=300 d=10	0,4	0,82	0,31	0,69	0,22	0,83	0,67	0,34
KNN_ROS RF n=300 d=10	0,4	0,82	0,31	0,69	0,22	0,83	0,67	0,34
KNN_Over2 RF n=200 d=10	0,4	0,82	0,31	0,69	0,22	0,83	0,67	0,34
Med_Over2 RF n=200 d=8	0,45	0,83	0,33	0,70	0,23	0,83	0,69	0,35
Med_Over2 RF n=300 d=8	0,45	0,83	0,33	0,70	0,23	0,83	0,69	0,36
Med_ROS RF n=300 d=8	0,45	0,83	0,33	0,70	0,23	0,82	0,69	0,35
Med_ROS RF n=200 d=8	0,45	0,83	0,33	0,70	0,22	0,82	0,69	0,35
Median RF n=300 d=10	0,4	0,83	0,32	0,71	0,23	0,82	0,70	0,36
Median RF n=200 d=10	0,4	0,83	0,32	0,71	0,23	0,81	0,70	0,36
KNN_Over2 RF n=300 d=8	0,45	0,82	0,31	0,70	0,22	0,81	0,69	0,35
KNN_Over2 RF n=200 d=8	0,45	0,82	0,31	0,70	0,22	0,81	0,69	0,35
KNN_ROS RF n=300 d=8	0,45	0,82	0,31	0,70	0,22	0,81	0,68	0,35
KNN_ROS RF n=200 d=8	0,45	0,82	0,31	0,70	0,22	0,81	0,68	0,35
Med_US RF n=200 d=12	0,5	0,84	0,35	0,74	0,25	0,80	0,73	0,38
Med_US RF n=300 d=12	0,5	0,84	0,35	0,74	0,25	0,80	0,73	0,38
Med_US RF n=200 d=10	0,5	0,84	0,35	0,74	0,25	0,80	0,73	0,38
Median RF n=200 d=8	0,45	0,83	0,32	0,72	0,24	0,80	0,72	0,36
Median RF n=300 d=8	0,45	0,83	0,33	0,72	0,23	0,80	0,71	0,36
Med_US RF n=300 d=10	0,5	0,83	0,34	0,74	0,24	0,80	0,73	0,37
Med_ROS RF n=300 d=12	0,4	0,83	0,33	0,73	0,24	0,80	0,73	0,37

3.2 Rendimiento de modelos XGB balanceados, filtrados por Sensibilidad > 0.80:

modelo	threshold	roc_auc	auprc	accuracy	precision	recall	specificity	f1
KNN_US XGB n=100 d=4	0,4	0,84	0,37	0,69	0,22	0,86	0,67	0,35
Med_US XGB n=50 d=6	0,4	0,84	0,38	0,69	0,22	0,86	0,67	0,35
Med_ROS XGB n=50 d=4	0,4	0,84	0,38	0,69	0,22	0,86	0,67	0,35
Med_US XGB n=50 d=4	0,4	0,84	0,37	0,68	0,22	0,86	0,66	0,34
Med_Over2 XGB n=50 d=4	0,4	0,84	0,38	0,69	0,22	0,86	0,67	0,35
KNN_US XGB n=50 d=4	0,4	0,83	0,36	0,67	0,21	0,86	0,65	0,34
KNN_US XGB n=50 d=6	0,4	0,84	0,36	0,68	0,22	0,86	0,66	0,35
Median XGB n=50 d=4	0,4	0,84	0,37	0,69	0,22	0,85	0,67	0,35
Med_US XGB n=100 d=6	0,4	0,85	0,39	0,70	0,23	0,85	0,68	0,36
KNN_ROS XGB n=50 d=4	0,4	0,84	0,37	0,68	0,22	0,85	0,66	0,34
KNN_US XGB n=100 d=6	0,4	0,84	0,37	0,69	0,22	0,85	0,67	0,35
Med_US XGB n=100 d=4	0,4	0,84	0,39	0,70	0,23	0,84	0,68	0,36
KNN XGB n=50 d=4	0,4	0,84	0,37	0,68	0,22	0,84	0,66	0,34
KNN_Over2 XGB n=50 d=4	0,4	0,84	0,37	0,68	0,22	0,84	0,66	0,34
Med_ROS XGB n=100 d=4	0,4	0,85	0,39	0,72	0,24	0,84	0,71	0,37
KNN_Over2 XGB n=100 d=4	0,4	0,84	0,38	0,71	0,23	0,84	0,70	0,36
Median XGB n=100 d=4	0,4	0,85	0,39	0,72	0,24	0,84	0,70	0,37
KNN XGB n=100 d=4	0,4	0,84	0,39	0,71	0,23	0,84	0,70	0,37
KNN_ROS XGB n=100 d=4	0,4	0,84	0,38	0,71	0,23	0,83	0,70	0,36
Med_Over2 XGB n=100 d=4	0,4	0,85	0,40	0,72	0,24	0,83	0,71	0,37
Med_Over2 XGB n=50 d=6	0,4	0,85	0,39	0,73	0,24	0,83	0,72	0,38
Med_ROS XGB n=50 d=6	0,4	0,84	0,37	0,73	0,24	0,82	0,72	0,37
KNN_US XGB n=50 d=4	0,45	0,83	0,36	0,71	0,23	0,82	0,70	0,36
Median XGB n=50 d=6	0,4	0,84	0,38	0,72	0,24	0,82	0,71	0,37
KNN_US XGB n=100 d=6	0,45	0,84	0,37	0,72	0,24	0,82	0,71	0,37
KNN_Over2 XGB n=50 d=6	0,4	0,83	0,37	0,72	0,23	0,82	0,71	0,36
Median XGB n=50 d=4	0,45	0,84	0,37	0,73	0,25	0,81	0,73	0,38
Med_US XGB n=50 d=4	0,45	0,84	0,37	0,72	0,24	0,81	0,71	0,37
KNN XGB n=50 d=6	0,4	0,83	0,37	0,72	0,23	0,81	0,71	0,36
Med_US XGB n=100 d=6	0,45	0,85	0,39	0,73	0,24	0,81	0,72	0,37
KNN_US XGB n=50 d=6	0,45	0,84	0,36	0,72	0,23	0,81	0,70	0,36
Med_US XGB n=100 d=4	0,45	0,84	0,39	0,73	0,24	0,81	0,72	0,37
Med_US XGB n=50 d=6	0,45	0,84	0,38	0,72	0,24	0,81	0,71	0,37
KNN_US XGB n=100 d=4	0,45	0,84	0,37	0,72	0,24	0,81	0,71	0,37
KNN_ROS XGB n=50 d=6	0,4	0,84	0,37	0,72	0,23	0,81	0,71	0,36
Med_ROS XGB n=50 d=4	0,45	0,84	0,38	0,74	0,25	0,81	0,73	0,38
KNN_ROS XGB n=50 d=4	0,45	0,84	0,37	0,72	0,24	0,81	0,72	0,37
KNN XGB n=50 d=4	0,45	0,84	0,37	0,72	0,24	0,80	0,72	0,37

KNN_Over2 XGB n=50 d=4	0,45	0,84	0,37	0,72	0,24	0,80	0,72	0,37
Med_Over2 XGB n=100 d=6	0,4	0,85	0,40	0,76	0,26	0,80	0,75	0,40
Med_Over2 XGB n=50 d=4	0,45	0,84	0,38	0,73	0,24	0,80	0,73	0,37
Med_ROS XGB n=100 d=4	0,45	0,85	0,39	0,76	0,26	0,80	0,75	0,39
KNN XGB n=100 d=4	0,45	0,84	0,39	0,75	0,25	0,80	0,74	0,39
Median XGB n=100 d=4	0,45	0,85	0,39	0,75	0,26	0,80	0,75	0,39
Med_Over2 XGB n=100 d=4	0,45	0,85	0,40	0,75	0,26	0,80	0,75	0,39

4.1 *Permutation Importance* de modelo RF elegido:

Indice	Variable	importance	std
36	h0_recuento_plaquetas	0,004740253	0,001153
13	codif_enf_renal_cronica	0,00400503	0,000838
31	h0_o2_flujo	0,002070233	0,001151
3	codif_insuficiencia_cardiaca	0,000677179	0,000698
39	h0_saturacion_o2_v	0,000522395	0,000321
27	h0_lactato_arterial	0,000367611	0,000296
25	h0_hemoglobina	0,000309568	0,000379
9	codif_enf_hematologica	0,000251524	9,87E-05
15	h0_albumina	0,000174132	0,000255
22	h0_fibrinogeno_calculado	0,000154784	0,000444
38	h0_saturacion_o2_a	0,000135436	0,000131
33	h0_potasio	7,73919E-05	0,000308
4	codif_epoc	-6,66134E-17	0,000293
7	codif_alzheimer	-1,9348E-05	3,87E-05
10	codif_enf_autoimmune	-1,9348E-05	3,87E-05
5	codif_asma	-5,80439E-05	7,74E-05
6	codif_demencia	-7,73919E-05	0,000248
0	sexo	-0,000154784	0,000452
14	h0_actividad_de_protrombina_porcentaje	-0,000174132	0,00108
43	h0_tas	-0,000212828	0,000927
48	h0_ph_v	-0,000232176	0,000404
44	h0_tiempo_de_tromboplastina_parcial_activado	-0,000251524	0,000264
29	h0_leucocitos	-0,000251524	0,000349
40	h0_spO2	-0,000270872	0,001397
46	h0_pco2_v	-0,000445003	0,000514
17	h0_bilirrubina_total	-0,000445003	0,00036
50	h0_po2_v	-0,000483699	0,00028
18	h0_co3h_a	-0,000522395	0,000284
30	h0_neutrofilos	-0,000561091	0,000394
21	h0_fc	-0,000677179	0,000346
35	h0_procalcitonina	-0,000677179	0,000392
49	h0_po2_a	-0,000696527	0,000155
8	codif_enf_oncologica	-0,000754571	0,000283
41	h0_t	-0,000773919	0,000478
37	h0_sodio	-0,000793267	0,00043
19	h0_co3h_v	-0,000986747	0,000598
2	codif_cardiopatía_isquemica	-0,001122182	0,00038
28	h0_lactato_venoso	-0,00123827	0,000654
32	h0_pco2_a	-0,001276966	0,000421
47	h0_ph_a	-0,001315662	0,000532
20	h0_creatinina	-0,00143175	0,001241

12	codif_diabetes	-0,001605882	0,001433
34	h0_proteina_c_reactiva	-0,001663926	9,48E-05
11	codif_hipertension	-0,001663926	0,000738
42	h0_tad	-0,002437845	0,000657
16	h0_alt_gpt	-0,00267002	0,00062
26	h0_linfocitos	-0,003134372	0,001447
23	h0_gammaglutamiltransferasa	-0,003927639	0,000468
24	h0_glucosa	-0,005746348	0,000686
45	h0_urea	-0,005785044	0,00111
1	edad_en_inicio_urgencias	-0,018322531	0,001945

4.2 Permutation importance de modelo XGB elegido.

Indice	Variable	importance	std
36	h0_recuento_plaquetas	0,009287027	0,0011
12	codif_diabetes	0,0028635	0,000556
40	h0_spO2	0,002786108	0,000552
21	h0_fc	0,001799362	0,000864
38	h0_saturacion_o2_a	0,00162523	0,000262
31	h0_o2_flujo	0,001586534	0,00072
8	codif_enf_oncologica	0,001102834	0,000413
2	codif_cardiopatía_isquemica	0,000851311	0,000495
19	h0_co3h_v	0,000831963	0,000413
17	h0_bilirrubina_total	0,000715875	0,000556
47	h0_ph_a	0,000715875	0,000365
22	h0_fibrinogeno_calculado	0,000619135	0,00039
18	h0_co3h_a	0,000619135	0,000355
49	h0_po2_a	0,000541743	0,000427
5	codif_asma	0,000503047	0,000255
43	h0_tas	0,000406307	0,000443
41	h0_t	0,000348264	0,000699
44	h0_tiempo_de_tromboplastina_parcial_activado	0,000328916	0,000291
3	codif_insuficiencia_cardiaca	0,000309568	0,000495
28	h0_lactato_venoso	0,000309568	0,000688
0	sexo	0,000309568	0,000542
50	h0_po2_v	0,000212828	0,000128
9	codif_enf_hematologica	0,000116088	0,000113
37	h0_sodio	0,000116088	0,000877
7	codif_alzheimer	9,67399E-05	0,000173
27	h0_lactato_arterial	-9,67399E-05	0,000252
10	codif_enf_autoinmune	-0,000116088	7,24E-05
6	codif_demencia	-0,000135436	9,87E-05
29	h0_leucocitos	-0,000135436	0,000775
4	codif_epoc	-0,000154784	0,000249
48	h0_ph_v	-0,000212828	0,000412
35	h0_procalcitonina	-0,000212828	0,000604
15	h0_albumina	-0,000270872	0,000215
33	h0_potasio	-0,000309568	0,000701
32	h0_pco2_a	-0,000386959	0,000237
39	h0_saturacion_o2_v	-0,000425655	0,000452
14	h0_actividad_de_protrombina_porcentaje	-0,000503047	0,000943
25	h0_hemoglobina	-0,000657831	0,00051
16	h0_alt_gpt	-0,001083487	0,001168
46	h0_pco2_v	-0,00114153	0,000128
30	h0_neutrofilos	-0,001489794	0,000972

42	h0_tad	-0,001547838	0,000474
34	h0_proteina_c_reactiva	-0,001741318	0,000942
11	codif_hipertension	-0,002302409	0,00043
20	h0_creatinina	-0,002708716	0,000497
23	h0_gammaglutamiltransferasa	-0,003482635	0,000648
24	h0_glucosa	-0,00400503	0,000569
26	h0_linfocitos	-0,004043726	0,001592
13	codif_enf_renal_cronica	-0,004933733	0,000639
45	h0_urea	-0,006307439	0,001193
1	edad_en_inicio_urgencias	-0,007816581	0,001533