Facturación y trabajo de la minipyme THUNDERCATS S.A.

Integrantes

Macarena Fernandez Urquiza
Arturo Lewinger
Federico Moreno
Sofia Perini
Bianca Picchetti

Contenido

1. *Paradigmas*
   - *Tierra-1*
   - *Tierra-2*
   - *Tierra-3*
2. *Metodología final*
3. *Modelos ganadores*
4. *Evaluación final*

1. Paradigmas

Y la importancia de la organización

Se atravesaron tres paradigmas

Con un objetivo común: realizar votación de los mejores modelos

TIERRA-1

Se puso el foco en la creación de variables. Se crearon cientos de ellas y se realizaron diversos experimentos.
Sin éxito.

TIERRA-2

Se vuelve a los datasets originales (sin variables creadas por nosotros). Se agregan predicciones de otros modelos como covariables. Se experimenta con los hiperparámetros.
Éxito moderado.

TIERRA-3

Se combina lo estudiado en Tierra-1 y Tierra-2. Se corrige comportamiento anómalo en el dataset original. Se agregan predicciones de otros modelos como covariables.
Éxito rotundo.

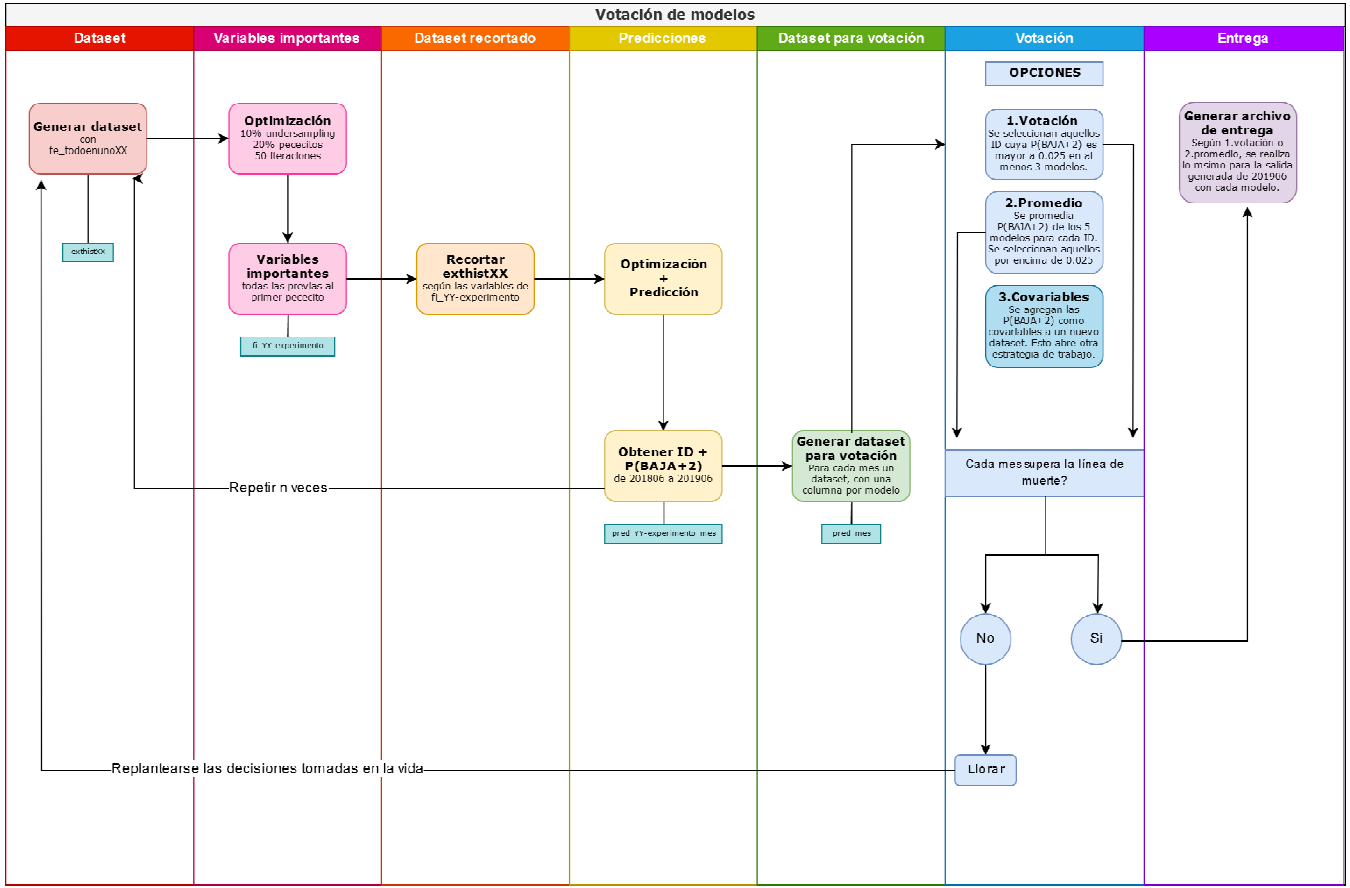


Diagrama de trabajo inicial

1. 1. Tierra-1

O cómo dejar volar la imaginación

Primeras 100 horas

Exploración
28 horas

Reuniones semanales durantes las primeras 5 semanas para discutir el dataset y realizar las tareas. Se crean las cuentas en la nube y se realizan los experimentos sin objetivo fijo.

Variables nuevas
80 horas

Agrupación de conceptos:
Tarjetas
Prestamos y adelantos
Tipo de cliente
Tipo de operaciones
Rentabilidad al banco
Marketing

Algunos ejemplos de variables creadas

Conceptos unificados

fecantprest := rowSums(cbind(cprestamos_personales, cprestamos_prendarios, cprestamos_hipotecarios) , na.rm=TRUE
fettdeuda := rowSums(cbind(mv_tadelantosefectivo, mprestamos_personales) , na.rm=TRUE
fettahorro := rowSums(cbind(mcaja_ahorro_Paquete, mcaja_ahorro_Nopaquete) , na.rm=TRUE
fettpatrimonio := rowSums(cbind(fettahorro, fettdeuda) , na.rm=TRUE)]

Ratios

fm_endeudamiento := mprestamos_personales / mv_mlimitecompra]
fm_pagominimo := mv_mpagominimo / fettdeuda]
fm_comisionel := mcomisiones / mrentabilidad]
fettdeuda_rango := fettdeuda/(fettdeuda__max - fettdeuda__min)]
fettdeuda_rango := fettdeuda/(fettdeuda__max - fettdeuda__min)]

Chiches

ECM_renta_avg := (mrentabilidad - mrentabilidad__avg)^2]
ECM_saldo_avg := (mcuentas_saldo - mcuentas_saldo__avg)^2]
ECM_sueldo_avg := (mplan_sueldo - mplan_sueldo__avg)^2]
ECM_ctactepaq_avg := (mcuenta_corriente_Paquete - mcuenta_corriente_Paquete__avg)^2]

Campeonato mundial de variables

"Que compitan ellas!"

01

MODELOS 1

Correr docenas de modelos con: Exthist + 200+ variables nuevas creadas

02

IMPORTANCIA

Observación de importancia de variables de cada modelo

03

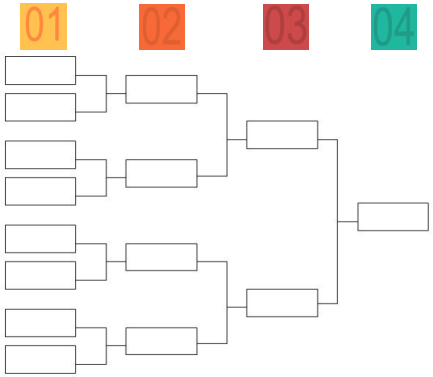
AGRUPACIÓN

Agrupación de variables que más se repetían y que, en promedio, tenian gain superior a un umbral

04

MODELOS 2

Correr el mismo algoritmo, ahora con la selección



Campeonato mundial de variables

"Que compitan ellas!"

01

MODELOS 1

Correr docenas de modelos con: Exthist + 200+ variables nuevas creadas

02

IMPORTANCIA

Observación de importancia de variables de cada modelo

03

AGRUPACIÓN

Agrupación de variables que más se repetían y que, en promedio, tenian gain superior a un umbral

04

MODELOS 2

Correr el mismo algoritmo, ahora con la selección

*Ganancia **menor** que con otros dataset*

1. 2. Tierra-2

O cómo volver a las bases

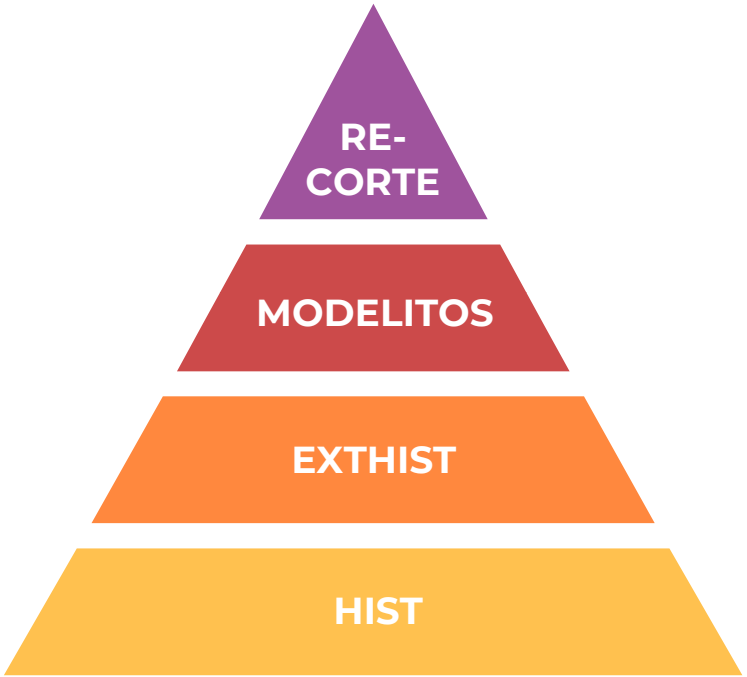
Se realizan experimentos de forma controlada comparando contra la línea de muerte

RECORTE DEL DATASET
Se conservan sólo las variables que se ubican en importancia por encima del primer pajarito

COVARIABLES
Se agregan como covariables las predicciones de otros modelos

VARIABLES NUEVAS + HISTORICAS
Nuevas variables (unificación de tarjetas) + ventana de regresión de 6 meses

VARIABLES HISTORICAS
Ventana de regresión de 6 meses



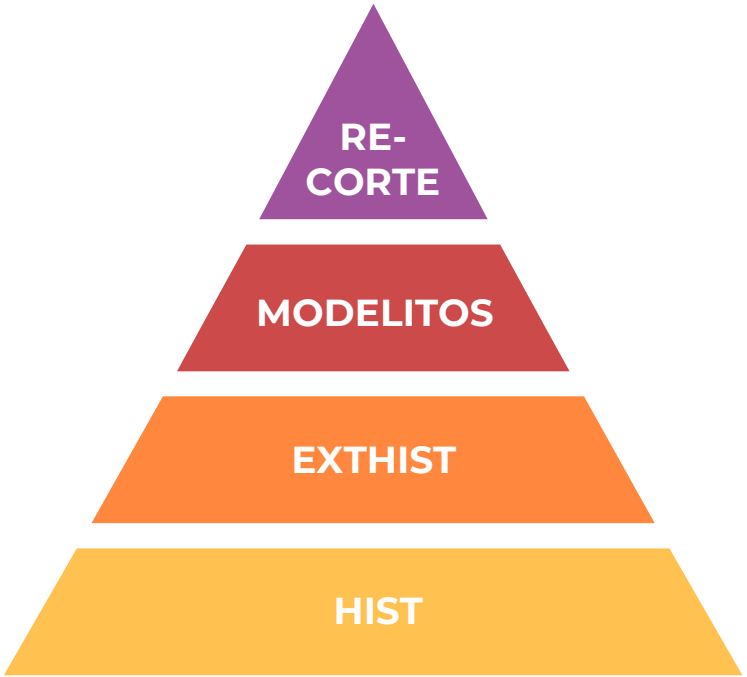
Se realizan experimentos de forma controlada comparando contra la línea de muerte

RECORTE DEL DATASET
Se conservan sólo las variables que se ubican en importancia por encima del primer pajarito

COVARIABLES
Se agregan como covariables las predicciones de otros modelos

VARIABLES NUEVAS + HISTORICAS
Nuevas variables (unificación de tarjetas) + ventana de regresión de 6 meses

VARIABLES HISTORICAS
Ventana de regresión de 6 meses



*Se supera la línea de muerte en **9 de 11 meses**, con diferencias **no mayores a 100k** en Abril*

1. 3. Tierra-3

O cómo poner lo mejor de cada uno

Los 3 pilares de Tierra-3

Aplicando lo aprendido en Tierra-1 y Tierra-2

LIMPIEZA

Se quitan del dataset original las variables con valores constantes durante todo el periodo.
Se introducen NA en aquellas variables con comportamiento atípico (sólo en el mes anómalo).

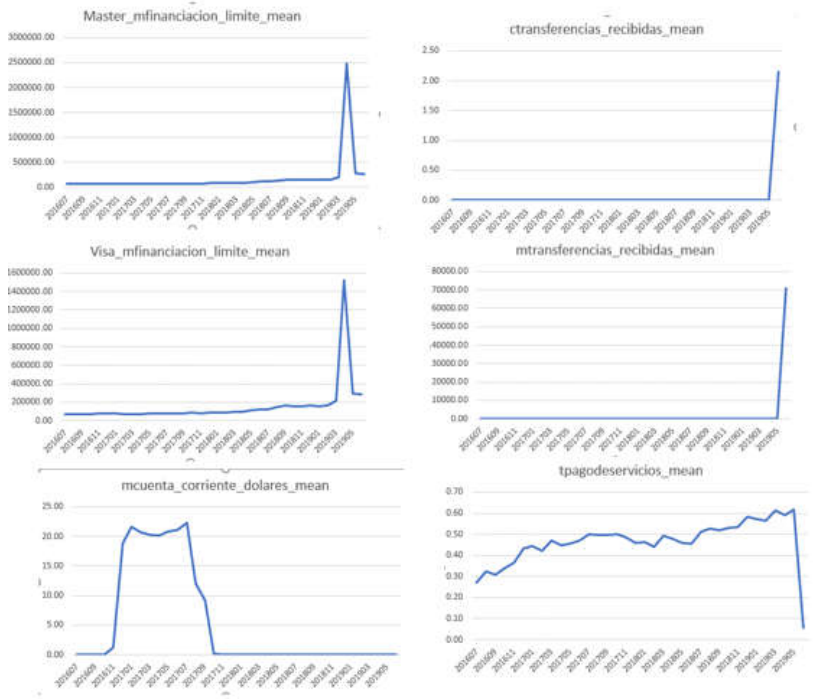
VARIABLES NUEVAS

Se agregan aquellas variables importantes para la clasificación.

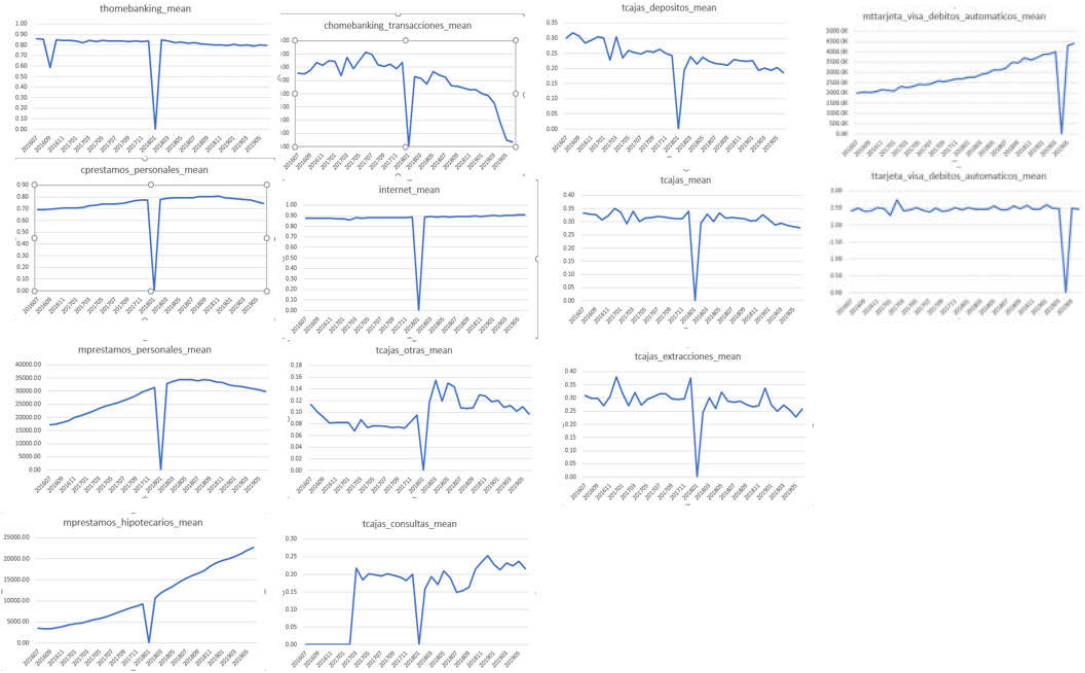
COVARIABLES

Se agregan como covariables las predicciones de otros modelos (no necesariamente consistentemente exitosos).

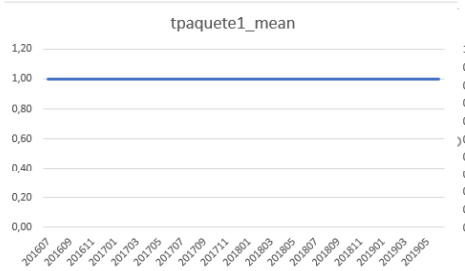
Variables anómalas

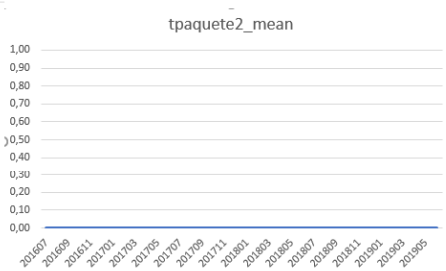


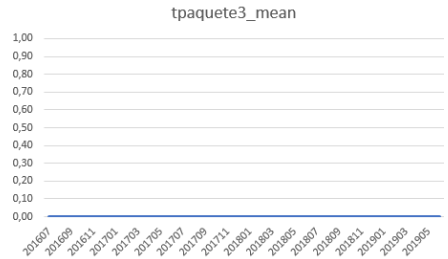
Variable-mes con media nula



Variables constantes







Se generan nuevos datasets

seleccionando entre distintas combinaciones

EXTHIST + VARIABLES NUEVAS

- Dataset original
- Variables nuevas
- Variables históricas

EXTHIST LIMPIO + VARIABLES NUEVAS

- Dataset original limpio
- Variables nuevas
- Variables históricas

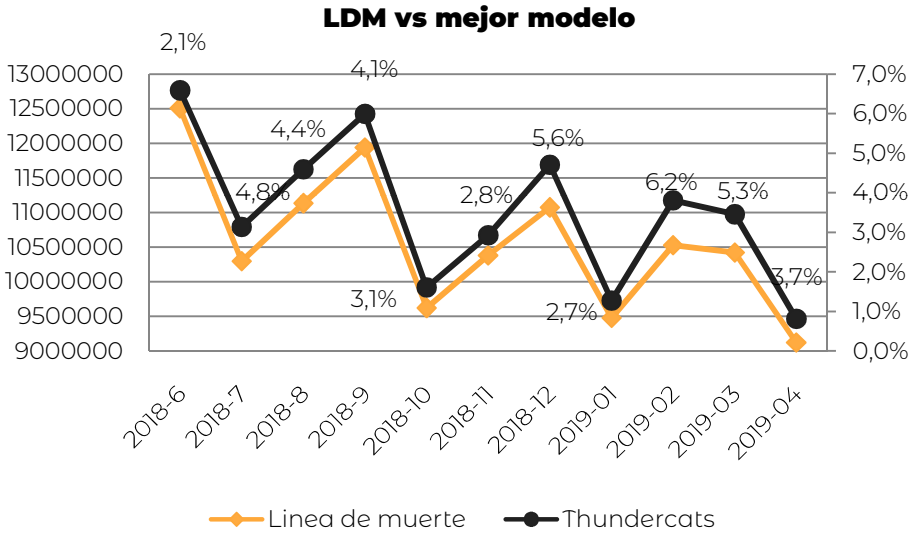
EXTHIST LIMPIO + CATEGORICAS

- Dataset original limpio
- Variables nuevas sólo categóricas
- Variables históricas

Ventana de regresión: 3 , 6 y/ó 12 meses

+ Probabilidades de otros modelos como covariables

Y todo fue bello



*Se supera la línea de muerte entre **9 y 11** (de 11 meses) con diferencias **mayores a 200k** en Abril*

3. Metodología final

O cómo dejar de sufrir

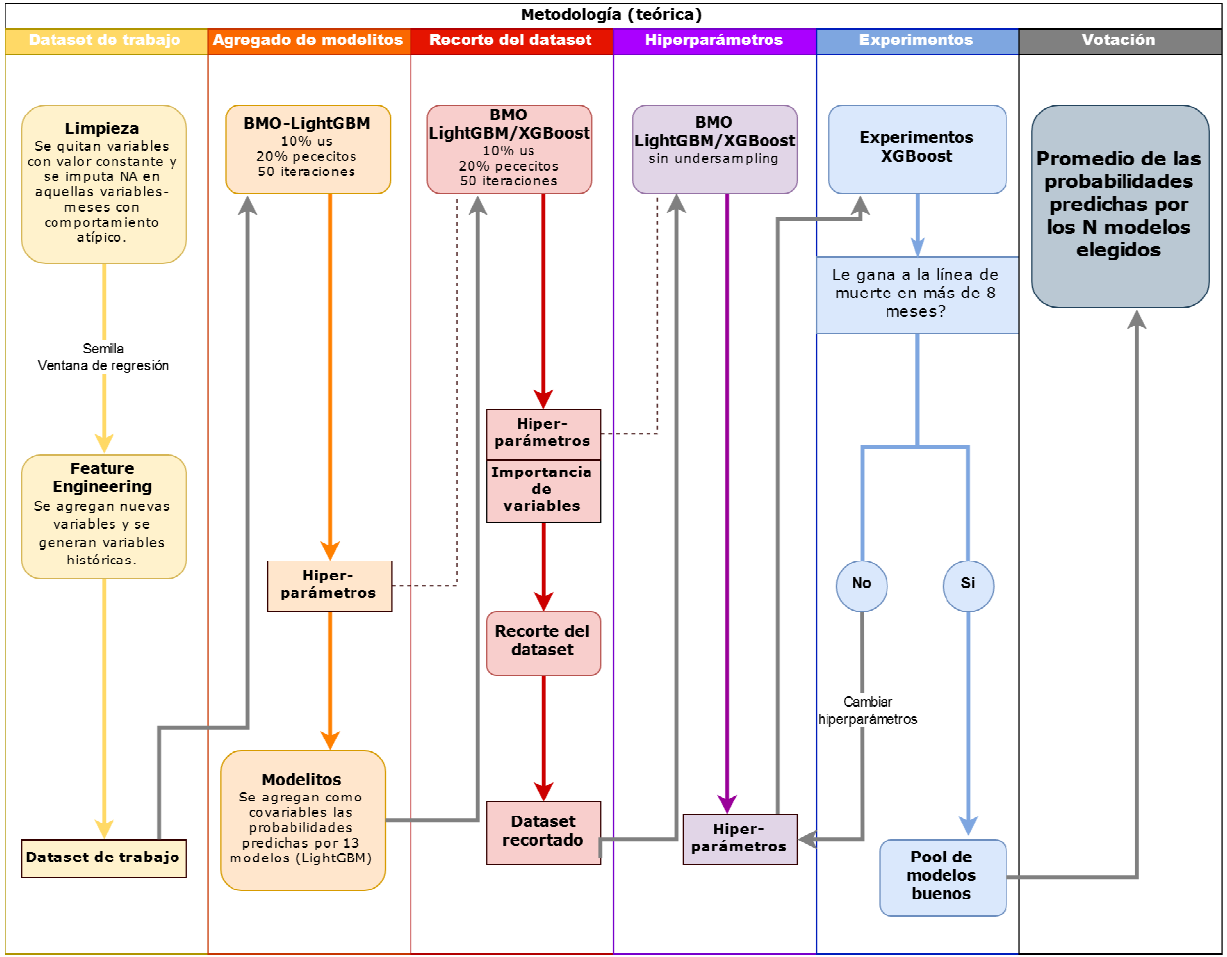


Diagrama de trabajo –teórico- final.

El último fin de semana

VIERNES
Definición del nuevo paradigma y metodología

SABADO
Crisis

DOMINGO
Todo empieza a funcionar

LUNES
Ya fue todo

4. Modelos ganadores

O cómo votar bien

Primer ensamble

Fede y Artur

	LINEA DE MUERTE	MODELO 1	MODELO 2	MODELO 3
DATASET	HIST	EXTHIST	EXTHIST LIMPIO (FE)	EXTHIST LIMPIO (FE)
VENTANA	6	3	3	6
VARIABLES	500+	2000+	1000+	900+
OPTIMIZACION		LGBM (SIN US)	LGBM (SIN US)	LGBM (10% US)
RECORTE		111	173	No
COVARIABLES		39 MODELITOS	13 MODELITOS	26 MODELITOS
PREDICCIÓN	XGBOOST	XGBOOST	XGBOOST	XGBOOST
PARÁMETROS	LDM	LDM	OPTIMIZADOS	OPTIMIZADOS

PROBABILIDAD DE CORTE: 0.027

Segundo ensamble

Maca

	LINEA DE MUERTE	MODELO 1	MODELO 4	MODELO 3
DATASET	HIST	EXTHIST	EXTHIST (CAT)	EXTHIST LIMPIO (FE)
VENTANA	6	3	6 Y 12	6
VARIABLES	500+	2000+	1000+	900+
OPTIMIZACION		LGBM (SIN US)	No	LGBM (10% US)
RECORTE		111	122	No
COVARIABLES		39 MODELITOS	13 MODELITOS	26 MODELITOS
PREDICCIÓN	XGBOOST	XGBOOST	XGBOOST	XGBOOST
PARÁMETROS	LDM	LDM	LDM	OPTIMIZADOS

PROBABILIDAD DE CORTE: 0.025

5. Evaluación final

O cómo triunfar (?)

400+

experimentos

Tierra-1

Fecha	N-	Experimento	Código original	Código editado	Dataset	Undersampling	Pececitos	Quien	Semilla modelo	Algoritmo
-	0	Linea de muerte	lineademuerte_UBA.r	-	paquete_premium_hist	No	No	GD	102191	XGBoost
11/11/2019	01	Argentine	lineademuerte_UBA.r	01-Argentine.r	paquete_premium_exthist	No	No	Biam!	102191	XGBoost
12/11/2019	02	Banded Sugar	lightgbm_directo_wfv_hist.r	02-Banded-sugar.r	paquete_premium_exthist	10%	136 - 16%	Biam!	102191	LightGBM
11/11/2019	03	Black Garden	lightgbm_directo_wfv_hist.r	03-Black-garden.r	paquete_premium_exthist	10%	160 - 19%	Artur	102191	LightGBM
12/11/2019	04	Bullet	lightgbm_directo_wfv_hist.r	04-Bullet.r	paquete_premium_exthist01	10%	136 - 16%	Sofi	188317	LightGBM
15/11/2019	05	Carpenter	lightgbm_directo_wfv_hist.r	05-Carpenter.r	paquete_premium_exthist03	10%	217 - 20%	Biam!	102191	LightGBM
16/11/2019	06	Electric	lightgbm_directo_wfv_auto.r	06-Electric.r	paquete_premium_exthist03	10%	217 - 20%	Biam!	102191	LightGBM
16/11/2019	07	False Honeypot	lightgbm_directo_wfv_baja.r	07-False-honeypot.r	paquete_premium_exthist04	10%	100 - %	Artur	102191	LightGBM
16/11/2019	08	Fire	lightgbm_directo_wfv_hist.r	08-Fire.r	paquete_premium_exthist05	10%	136 - %	Sofi	188317	LightGBM
17/11/2019	09	Pavement	lineademuerte_UBA.r		paquete_premium_exthist06	No	No	Fede	102191	XGBoost
17/11/2019	10	Saint Valentine	lightgbm_directo_wfv_hist.r	10-saint-valentine.r	fi_exthist05	No	30 - 20%	Sofi	188317	LightGBM
18/11/2019	11	Pharaoh	lightgbm_directo_wfv_hist.r	11-Pharaoh.r	paquete_premium_exthist08	10%	136 - 6%	Maca	661259	LightGBM
18/11/2019	12	Red Harvester	lightgbm_directo_wfv_hist.r	12-Red-harvester.r	paquete_premium_exthist06	10%	136 - %	Fede	188317	LightGBM
19/11/2019	13	Cocktail	lightgbm_directo_wfv_hist.r	13-Cocktail.r	fi_exthist08	10%	20%	Maca	661259	LightGBM
19/11/2019	14	Acrobat	lightgbm_directo_wfv_hist.r	14-Acrobat.r	fi_exthist10	10%	136 - %	Fede	188317	LightGBM
19/11/2019	15	Yellow Crazy	lightgbm_directo_wfv_hist.r	15-Yellow-crazy.r	paquete_premium_exthist09	10%	281 - 20%	Sofi	188317	LightGBM
20/11/2019	16	Black Carpenter	lightgbm_directo_wfv_hist.r	15-black-carpenter.r	paquete_premium_exthist11	10%	281 - 20%	Biam!	102191	LightGBM
20/11/2019	17	Dracula						Artur		LightGBM
20/11/2019	18	Azteca						Artur		LightGBM
20/11/2019	19	Goblin						Artur		LightGBM
20/11/2019	20	Longhorn crazy	lightgbm_directo_wfv_hist.r	20-longhorn-crazy.r	fi_exthist11	10%	20 - 20%	Biam!	102191	LightGBM
21/11/2019	21	Wood	xgboost_directo_wfv.r	21-Wood.r	fi_exthist09	10%		Sofi	102191	XGBoost
22/11/2019	22	Mound	xgboost_SOFI_revisadoGD.R	22-Mound.r	paquete_premium_hist	10%		GD	102192	XGBoost

Tierra-1

Fecha	N-	Experimento	Código original	Código editado	Dataset	Undersampling	Pececitos	Quien	Semilla modelo	Algoritmo
-	0	Linea de muerte	lineademuerte_UBA.r	-	paquete_premium_hist	No	No	GD	102191	XGBoost

rediccio	rain_desdrain_hast	inea de muer	rain_desdrain_hast	Argentinc	rain_desdrain_hast	landed Sagg	rain_desdrain_hast	Black Garde	rain_desdrain_hast	Bullet	rain_desdrain_hast	Carpenter	rain_desdrain_hast	Electric	rain_desdrain_hast	alsc Honeyp	rain_desdrain_hast	Fire
201904	201805	201902	201805	201902	9,048,000	201805	201902	8,531,500	201805	201902	8,576,500	201805	201902	8,089,500	201805	201902	8,853,000	201806	201902	4,433,000	8,862,000	201805	201902	9,051,500
201903	201804	201901	201804	201901	10,465,500				201806	201901	8,673,500	201805	201901	8,866,500										
201902	201803	201812	201803	201812	10,675,500				201804	201812	10,643,500	201804	201812	9,048,500										
201901	201802	201811	201802	201811					201802	201811	9,288,000	201802	201811	8,678,500										
201812	201801	201810	201801	201810					201801	201810	10,372,500	201801	201810	9,515,500										
201811	201712	201803	201712	201803					201712	201803	10,063,500	201802	201803	8,722,500										
201810	201711	201808	201711	201808					201711	201808	9,332,000	201711	201808	8,285,500										
201809	201710	201807	201710	201807					201711	201807	11,672,500	201711	201807	10,727,500										
201808	201709	201806	201709	201806					201711	201806	10,344,500	201709	201806	8,789,500										
201807	201708	201805	201708	201805					201708	201805	10,081,500	201708	201805	9,612,000										
201806	201707	201804	201707	201804					201708	201804	12,005,500	201708	201804	11,277,000										

rediccio	rain_desdrain_hast	inea de muer	rain_desdrain_hast	Parcmeat	rain_desdrain_hast	aint Yalcatin	rain_desdrain_hast	int Yalcatinc (U	rain_desdrain_hast	Pharaoh	rain_desdrain_hast	Red Harvester	rain_desdrain_hast	Cocktail	rain_desdrain_hast	Acrobat	rain_desdrain_hast	ellow Cra
201904	201805	201902	201805	201902	9,122,500	8,403,000	201806	201902	9,268,500	201806	201902	9,112,000	201805	201902	8,875,000	201805	201902	9,132,500	201805	201902	8,842,500	201805	201902	8,823,000	201805	201902	8,444,500
201903	201804	201901			10,423,000																	201806	201901	10,042,000			
201902	201803	201812			10,528,000																	201806	201812	10,489,500			
201901	201802	201811			3,473,000																						
201812	201801	201810			11,073,000																						
201811	201712	201803			10,382,500																						
201810	201711	201808			9,620,500																						
201809	201710	201807			11,939,500																						
201808	201709	201806			11,132,500																						
201807	201708	201805			10,298,000																						
201806	201707	201804			12,505,500																						

rediccio	rain_desdrain_hast	inea de muer	rain_desdrain_hast	lack Carpent	rain_desdrain_hast	Dracula	rain_desdrain_hast	Azteca	rain_desdrain_hast	Goblin	rain_desdrain_hast	Longhorn crazy	rain_desdrain_hast	Wood	rain_desdrain_hast	Mound	rain_desdrain_hasta
201904	201805	201902	201806	201902	8,894,500						201805	201902	8,867,000	201803	201902	1,977,500	201803	201902	8,839,000	
201903	201804	201901	201805	201901	10,388,500						201806	201901	9,482,500	201809	201901	2,019,000	201802	201901	10,553,000	
201902	201803	201812			10,528,000						201803	201812	10,626,500							
201901	201802	201811			3,473,000						201802	201811	8,744,000							
201812	201801	201810			11,073,000						201803	201810	10,631,500							
201811	201712	201803			10,382,500						201712	201803	10,306,500							
201810	201711	201808			9,620,500															
201809	201710	201807			11,939,500															
201808	201709	201806			11,132,500															
201807	201708	201805			10,298,000															
201806	201707	201804			12,505,500															

20/11/2019	20	Longhorn crazy	lightgbm_directo_wfv_hist.r	20-longhorn-crazy.r	fi_exthist11	10%	20 - 20%	Biam!	102191	LightGBM
21/11/2019	21	Wood	xgboost_directo_wfv.r	21-Wood.r	fi_exthist09	10%		Sofi	102191	XGBoost
22/11/2019	22	Mound	xgboost_SOFI_revisadoGD.R	22-Mound.r	paquete_premium_hist	10%		GD	102192	XGBoost

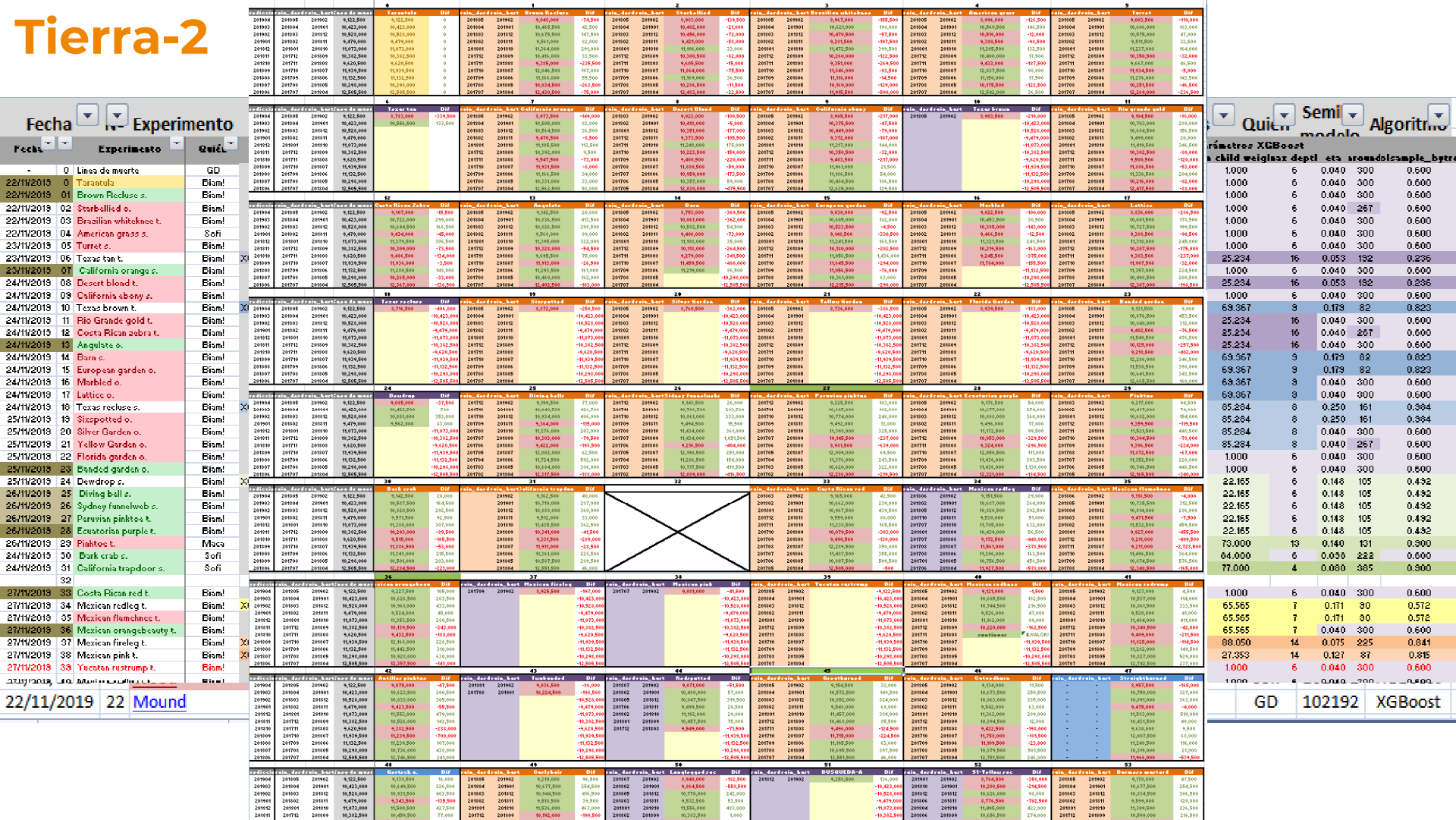
Tierra-2

Fecha	N° Experimento	Código original	Código editado	Dataset	Undersampling	Pececitos	Quién	Semi modelo	Algoritmo

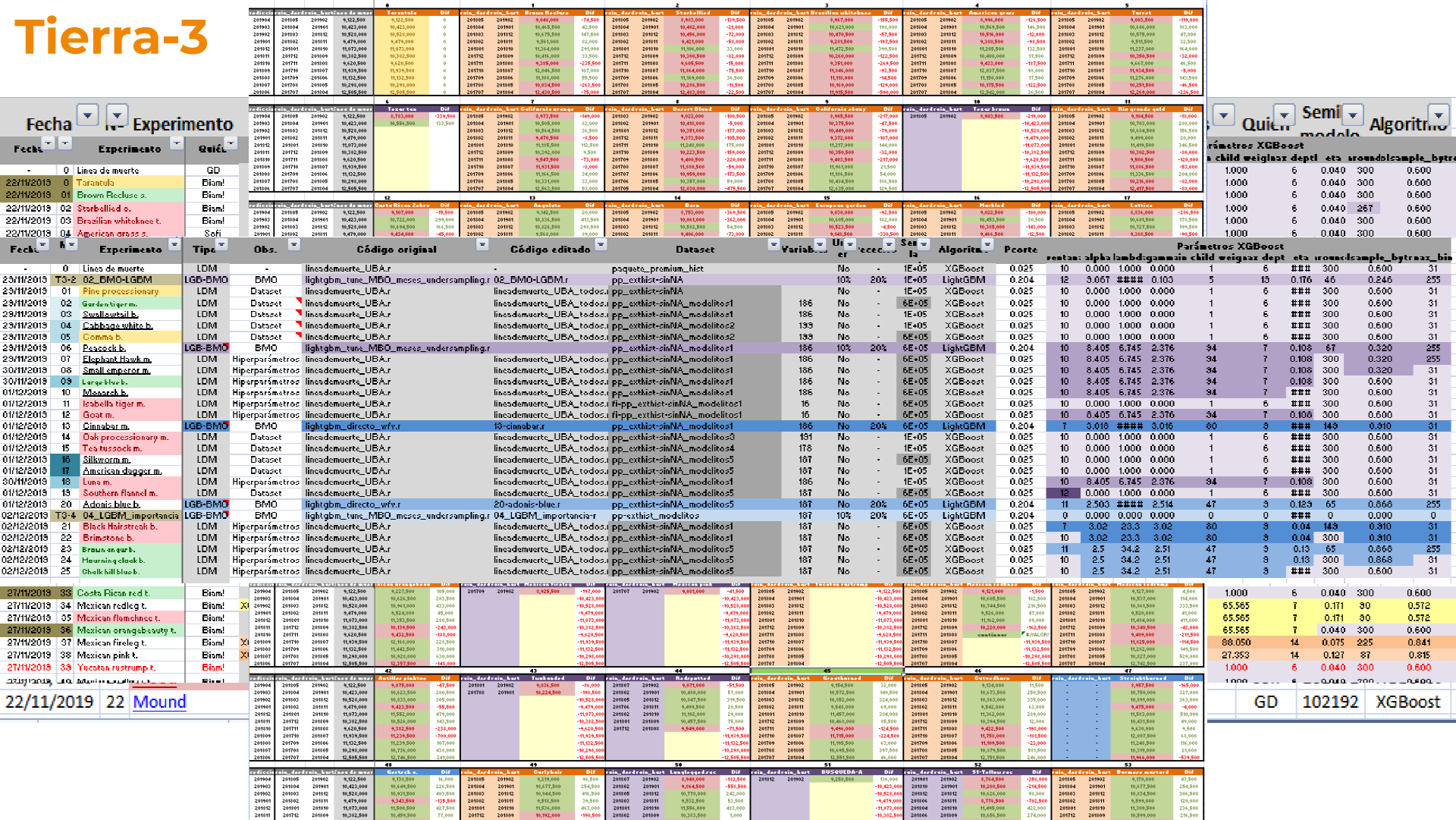
Fecha	Experimento	Quién	Tipo	Obs.	Código original	Código editado	Dataset	Undersampling	Pececitos	Semilla	Algoritmo	Pcorte	Parámetros XGBoost								
													nrestimators	alpha	lambda	gamma	min child weight	max depth	eta	nrounds	colsample_bytree
-	0 Lines de muerte	GD	LDM	-	lineademuerte_UBA.r	-	paquete_premium_hist	No	-	1E+05	XGBoost	0.025	10	0.000	1.000	0.000	1.000	6	0.040	300	0.600
22/11/2019	0 Tarantula	Biam!	LDM	Test	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_hist	No	-	1E+05	XGBoost	0.025	10	0.000	1.000	0.000	1.000	6	0.040	300	0.600
22/11/2019	01 Brown Recluse s.	Biam!	LDM	Dataset	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_exthist	No	-	1E+05	XGBoost	0.025	10	0.000	1.000	0.000	1.000	6	0.040	300	0.600
22/11/2019	02 Starbellied o.	Biam!	LDM	267	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_hist	No	-	1E+05	XGBoost	0.025	10	0.000	1.000	0.000	1.000	6	0.040	267	0.600
22/11/2019	03 Brazilian whiteknee t.	Biam!	LDM	us 10%	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_hist	10%	-	1E+05	XGBoost	0.204	10	0.000	1.000	0.000	1.000	6	0.040	300	0.600
22/11/2019	04 American grass s.	Sofi	LDM	semilla	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_hist	No	-		XGBoost	0.025	10	0.000	1.000	0.000	1.000	6	0.040	300	0.600
23/11/2019	05 Turret s.	Biam!	LDM	us 30%	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_hist	30%	-	1E+05	XGBoost	0.204	10	0.000	1.000	0.000	1.000	6	0.040	300	0.600
23/11/2019	06 Texas tan t.	Biam!	XGB-BMO	BMO		xgboost_directo_wfv.r	paquete_premium_hist	0.3	0.2	1E+05	XGBoost	0.079	12	3.993	35.070	8.058	25.234	16	0.053	192	0.236
23/11/2019	07 California orange s.	Biam!	LDM	us 50%	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_hist	50%	-	1E+05	XGBoost	0.043	10	0.000	1.000	0.000	1.000	6	0.040	300	0.600
24/11/2019	08 Desert blond t.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_hist	No	-	1E+05	XGBoost	0.025	12	3.993	35.070	8.058	25.234	16	0.053	192	0.236
24/11/2019	09 California ebony s.	Biam!	LDM	Dataset	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	fi_hist1	No	-	1E+05	XGBoost	0.025	10	0.000	1.000	0.000	1.000	6	0.040	300	0.600
24/11/2019	10 Texas brown t.	Biam!	XGB-BMO	BMO		xgboost_directo_wfv.r	paquete_premium_exthist	0.1	0.2	1E+05	XGBoost	0.204	10	12.001	14.339	5.687	69.367	3	0.179	82	0.823
24/11/2019	11 Rio Grande gold t.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_hist	No	-	1E+05	XGBoost	0.025	10	3.993	35.070	8.058	25.234	16	0.040	300	0.600
24/11/2019	12 Costa Rican zebra t.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_hist	No	-	1E+05	XGBoost	0.025	10	3.993	35.070	8.058	25.234	16	0.040	267	0.600
24/11/2019	13 Angulate o.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_exthist	No	-	1E+05	XGBoost	0.025	10	3.993	35.070	8.058	25.234	16	0.040	300	0.600
24/11/2019	14 Barn o.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	fi_exthist01-T2	No	-	1E+05	XGBoost	0.025	10	12.001	14.339	5.687	69.367	3	0.179	82	0.823
24/11/2019	15 European garden o.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_exthist	No	-	1E+05	XGBoost	0.025	10	12.001	14.339	5.687	69.367	3	0.179	82	0.823
24/11/2019	16 Marbled o.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	fi_exthist01-T2	No	-	1E+05	XGBoost	0.025	10	12.001	14.339	5.687	69.367	3	0.040	300	0.600
24/11/2019	17 Lattice o.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_exthist	No	-	1E+05	XGBoost	0.025	10	12.001	14.339	5.687	69.367	3	0.040	300	0.600
24/11/2019	18 Texas recluse s.	Biam!	XGB-BMO	BMO		xgboost_directo_wfv.r	paquete_premium_exthist	No	0.2	1E+05	XGBoost	0.025	10	23.828	10.787	3.620	85.284	8	0.250	161	0.364
25/11/2019	19 Sixspotted o.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_exthist	No	-	1E+05	XGBoost	0.025	10	23.828	10.787	3.620	85.284	8	0.250	161	0.364
25/11/2019	20 Silver Garden o.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_exthist	No	-	1E+05	XGBoost	0.025	10	23.828	10.787	3.620	85.284	8	0.040	300	0.600
25/11/2019	21 Yellow Garden o.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_exthist	No	-	1E+05	XGBoost	0.025	10	23.828	10.787	3.620	85.284	8	0.040	267	0.600
25/11/2019	22 Florida garden o.	Biam!	LDM	Dataset	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_exthist (fast)	No	-	1E+05	XGBoost	0.025	10	0.000	1.000	0.000	1.000	6	0.040	300	0.600
25/11/2019	23 Banded garden o.	Biam!	LDM	Dataset	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_exthist_modelitosxvsl	No	-	1E+05	XGBoost	0.025	10	0.000	1.000	0.000	1.000	6	0.040	300	0.600
25/11/2019	24 Dewdrop s.	Biam!	XGB-BMO	BMO		xgboost_directo_wfv.r	paquete_premium_exthist_modelitosxvsl	0.1	0.2	1E+05	XGBoost	0.204	15	12.343	40.022	7.738	22.165	6	0.148	105	0.492
26/11/2019	25 Diving bell s.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	paquete_premium_exthist_modelitosxvsl	No	-	1E+05	XGBoost	0.025	15	12.343	40.022	7.738	22.165	6	0.148	105	0.492
26/11/2019	26 Sydney funnelweb s.	Biam!	LDM	Dataset	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	fi_pomelitos01	No	-	1E+05	XGBoost	0.025	15	12.343	40.022	7.738	22.165	6	0.148	105	0.492
26/11/2019	27 Peruvian pinktoe t.	Biam!	LDM	Dataset	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	pomelitos_cov1	No	-	1E+05	XGBoost	0.025	15	12.343	40.022	7.738	22.165	6	0.148	105	0.492
26/11/2019	28 Ecuatorian purple t.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	pomelitos_cov1	No	-	1E+05	XGBoost	0.025	15	12.343	40.022	7.738	22.165	6	0.148	105	0.492
26/11/2019	29 Pinktoe t.	Maca	LDM	Dataset	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	exthist-flag	No	-	1E+05	XGBoost	0.025	12	26.000	10.000	2.000	73.000	13	0.140	131	0.900
24/11/2019	30 Bark crab s.	Sofi	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	fi_exthist05	No	-	1E+05	XGBoost	0.025	12	11.000	27.000	3.500	84.000	6	0.036	222	0.600
24/11/2019	31 California trapdoor s.	Sofi	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	fi_exthist05	No	-	1E+05	XGBoost	0.025	18	3.000	21.000	1.000	77.000	4	0.080	385	0.900
	32																				
27/11/2019	33 Costa Rican red t.	Biam!	LDM	Dataset	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	pomelitos_cov2	No	-	1E+05	XGBoost	0.204	10	0.000	1.000	0.000	1.000	6	0.040	300	0.600
27/11/2019	34 Mexican redleg t.	Biam!	XGB-BMO	BMO		xgboost_directo_wfv.r	pomelitos_cov2	0.1	0.2	1E+05	XGBoost	0.068	3	13.050	15.551	5.463	65.565	7	0.171	30	0.572
27/11/2019	35 Mexican flamehose t.	Biam!	LDM	Dataset	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	pomelitos_cov2	No	-	1E+05	XGBoost	0.025	3	13.050	15.551	5.463	65.565	7	0.171	30	0.572
27/11/2019	36 Mexican orangebeauty t.	Biam!	LDM	Hiperparámetros	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	pomelitos_cov2	No	-	1E+05	XGBoost	0.025	10	13.050	15.551	5.463	65.565	7	0.040	300	0.600
27/11/2019	37 Mexican fireleg t.	Biam!	XGB-BMO	BMO		xgboost_directo_wfv.r	exthist10_cov1	0.1	0.2	1E+05	XGBoost	0.204	18	17.528	19.033	8.436	88.050	14	0.075	225	0.841
27/11/2019	38 Mexican pink t.	Biam!	XGB-BMO	BMO		xgboost_directo_wfv.r	exthist13_cov1	0.1	0.2	1E+05	XGBoost	0.204	18	5.165	4.838	10.629	27.353	14	0.127	87	0.815
27/11/2019	39 Yucatan rustrump t.	Biam!	LDM	Dataset	lineademuerte_UBA.r	lineademuerte_UBA_todos.r	fi_exthist10_cov1	No	-	1E+05	XGBoost	0.025	10	0.000	1.000	0.000	1.000	6	0.040	300	0.600
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

22/11/2019	22	Mound	xgboost_SOFI_revisadoGD.R	22-Mound.r	paquete_premium_hist	10%		GD	102192	XGBoost

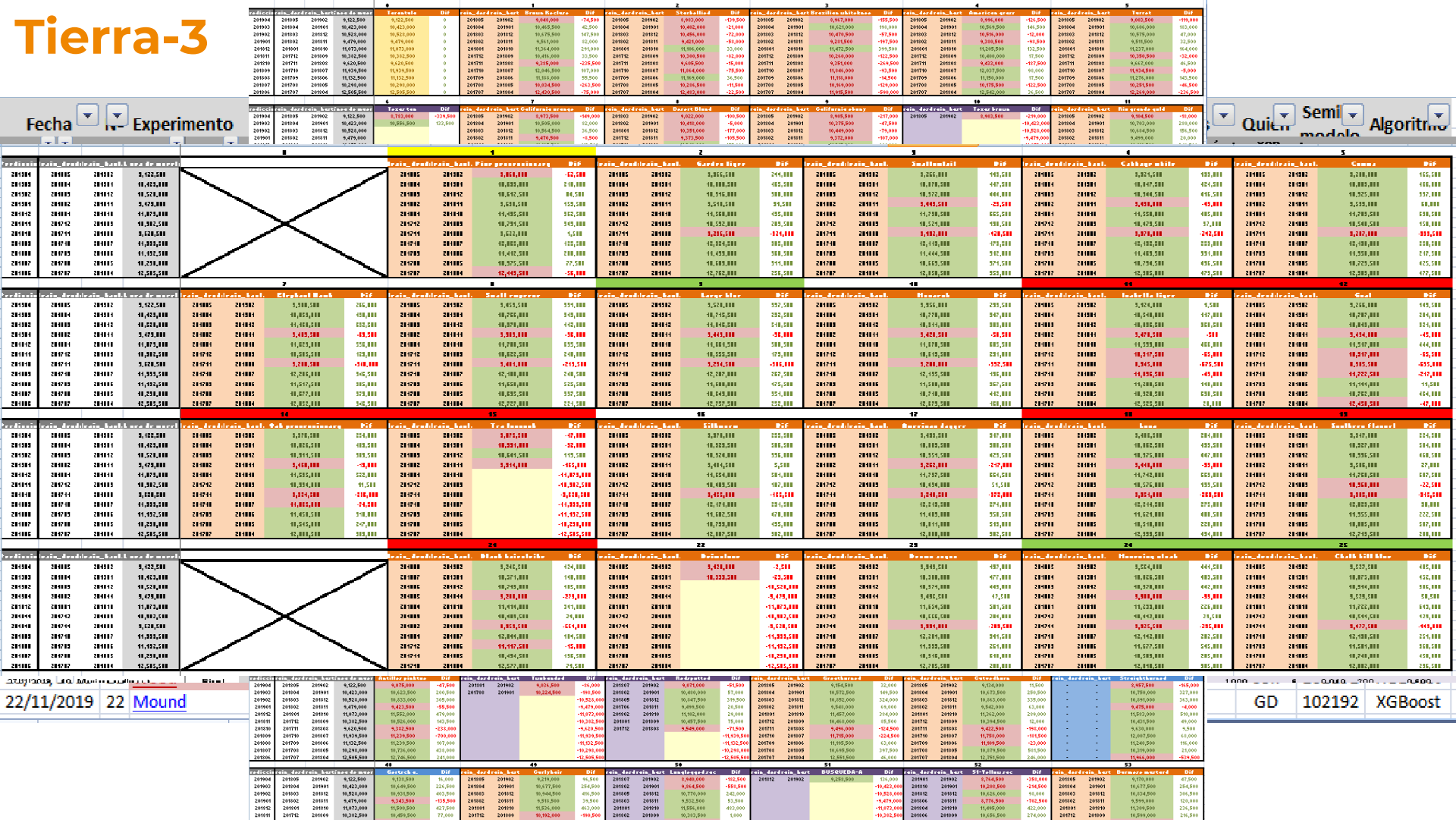
Tierra-2


Fecha
Experimento
Quién
Semi modelo
Algoritmo
GD
102192
XGBoost
22/11/2019
22
Mound

Tierra-3


Fecha
Experimento
Quién
Semi modelo
Algoritmo
Parámetros XGBoost
Tipo
Obs.
Código original
Código editado
Dataset
Variables
Recency
Semilla
Algoritmo
Pcorte
num_leaves
alpha
lambda
min_gain_to_split
max_depth
min_child_weight
bagging_fraction
bagging_freq
feature_fraction
scale_pos_weight
0 Linea de muerte GD
22/11/2019 01 Brown Recluse s. Biam!
22/11/2019 02 Starbellied s. Biam!
22/11/2019 03 Brazilian whiteknee t. Biam!
22/11/2019 04 American grass s. Sofi
23/11/2019 T3-2 02_BMO-LGBM LGB-BMO BMO lightgbm_tunc_MBO_meses_undersampling.r 02_BMO-LGBM.r pp_exthist-sinNA 10% 20% 1E+05 LightGBM 0.204 12 3.067 ##### 0.103 5 13 0.176 46 0.846 255
23/11/2019 01 Pine processionary LDM Dataset lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA No - 1E+05 XGBoost 0.025 10 0.000 1.000 0.000 1 6 ### 300 0.600 31
23/11/2019 02 Garden tiger m. LDM Dataset lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos1 186 No - 1E+05 XGBoost 0.025 10 0.000 1.000 0.000 1 6 ### 300 0.600 31
23/11/2019 03 Swallowtail b. LDM Dataset lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos1 186 No - 1E+05 XGBoost 0.025 10 0.000 1.000 0.000 1 6 ### 300 0.600 31
23/11/2019 04 Cabbage white b. LDM Dataset lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos2 186 No - 1E+05 XGBoost 0.025 10 0.000 1.000 0.000 1 6 ### 300 0.600 31
23/11/2019 05 Comma b. LDM Dataset lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos2 186 No - 6E+05 XGBoost 0.025 10 0.000 1.000 0.000 1 6 ### 300 0.600 31
23/11/2019 06 Peacock b. LGB-BMO BMO lightgbm_tunc_MBO_meses_undersampling.r pp_exthist-sinNA_modelitos1 186 10% 20% 6E+05 LightGBM 0.204 10 8.405 6.745 2.376 34 7 0.108 67 0.320 255
23/11/2019 07 Elephant Hawk m. LDM Hiperparámetros lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos1 186 No - 6E+05 XGBoost 0.025 10 8.405 6.745 2.376 34 7 0.108 300 0.320 255
30/11/2019 08 Small emperor m. LDM Hiperparámetros lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos1 186 No - 6E+05 XGBoost 0.025 10 8.405 6.745 2.376 34 7 0.108 300 0.320 31
30/11/2019 09 Large blue b. LDM Hiperparámetros lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos1 186 No - 6E+05 XGBoost 0.025 10 8.405 6.745 2.376 34 7 0.108 300 0.600 31
01/12/2019 10 Monarch b. LDM Hiperparámetros lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos1 186 No - 6E+05 XGBoost 0.025 10 8.405 6.745 2.376 34 7 ### 300 0.600 31
01/12/2019 11 Isabella tiger m. LDM Hiperparámetros lineademuerte_UBA.r lineademuerte_UBA_todos.r fi-pp_exthist-sinNA_modelitos1 16 No - 6E+05 XGBoost 0.025 10 0.000 1.000 0.000 1 6 ### 300 0.600 31
01/12/2019 12 Goat m. LDM Hiperparámetros lineademuerte_UBA.r lineademuerte_UBA_todos.r fi-pp_exthist-sinNA_modelitos1 16 No - 6E+05 XGBoost 0.025 10 8.405 6.745 2.376 34 7 0.108 300 0.600 31
01/12/2019 13 Cinnaber m. LGB-BMO BMO lightgbm_directo_wfv.r 13-cinnaber.r pp_exthist-sinNA_modelitos1 186 No 20% 6E+05 LightGBM 0.204 7 3.018 ##### 3.016 60 3 ### 143 0.310 31
01/12/2019 14 Oak processionary m. LDM Dataset lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos3 191 No - 1E+05 XGBoost 0.025 10 0.000 1.000 0.000 1 6 ### 300 0.600 31
01/12/2019 15 Tea tussock m. LDM Dataset lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos4 178 No - 1E+05 XGBoost 0.025 10 0.000 1.000 0.000 1 6 ### 300 0.600 31
01/12/2019 16 Silkworm m. LDM Dataset lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos5 187 No - 6E+05 XGBoost 0.025 10 0.000 1.000 0.000 1 6 ### 300 0.600 31
01/12/2019 17 American dagger m. LDM Dataset lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos5 187 No - 1E+05 XGBoost 0.025 10 0.000 1.000 0.000 1 6 ### 300 0.600 31
30/11/2019 18 Luna m. LDM Hiperparámetros lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos1 186 No - 1E+05 XGBoost 0.025 10 8.405 6.745 2.376 34 7 0.108 300 0.600 31
01/12/2019 19 Southern flannel m. LDM Dataset lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos5 187 No - 6E+05 XGBoost 0.025 12 0.000 1.000 0.000 1 6 ### 300 0.600 31
01/12/2019 20 Adonis blue b. LGB-BMO BMO lightgbm_directo_wfv.r 20-adonis-blue.r pp_exthist-sinNA_modelitos5 187 No 20% 6E+05 LightGBM 0.204 11 2.503 ##### 2.514 47 3 0.123 65 0.868 255
02/12/2019 T3-4 04_LGBM_importancia LGB-BMO BMO lightgbm_tunc_MBO_meses_undersampling.r 04_LGBM_importancia-r pp-exthist_modelitos 187 10% 20% 6E+05 LightGBM 0.204 0 0.000 0.000 0.000 0 0 ### 0 0.000 0
02/12/2019 21 Black Hairstreak b. LDM Hiperparámetros lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos1 187 No - 6E+05 XGBoost 0.025 7 3.02 23.3 3.02 80 3 0.04 143 0.310 31
02/12/2019 22 Brimstone b. LDM Hiperparámetros lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos1 187 No - 6E+05 XGBoost 0.025 10 3.02 23.3 3.02 80 3 0.04 300 0.310 31
02/12/2019 23 Brown argus b. LDM Hiperparámetros lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos5 187 No - 6E+05 XGBoost 0.025 11 2.5 34.2 2.51 47 3 0.13 65 0.868 255
02/12/2019 24 Mourning cloak b. LDM Hiperparámetros lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos5 187 No - 6E+05 XGBoost 0.025 10 2.5 34.2 2.51 47 3 0.13 300 0.868 31
02/12/2019 25 Chalk hill blue b. LDM Hiperparámetros lineademuerte_UBA.r lineademuerte_UBA_todos.r pp_exthist-sinNA_modelitos5 187 No - 6E+05 XGBoost 0.025 10 2.5 34.2 2.51 47 3 ### 300 0.600 31
27/11/2019 33 Costa Rican red t. Biam!
27/11/2019 34 Mexican redleg t. Biam!
27/11/2019 35 Mexican flamehance t. Biam!
27/11/2019 36 Mexican orangebeauty t. Biam!
27/11/2019 37 Mexican fireleg t. Biam!
27/11/2019 38 Mexican pink t. Biam!
27/11/2019 39 Yucatan rustrump t. Biam!
22/11/2019 22 Mound GD 102192 XGBoost

Tierra-3



Fecha | N° Experimento | Quién | Semi modelo | Algoritmo

22/11/2019 | 22 | Mound | GD | 102192 | XGBoost

5

Personas **colaborando**

9

Cuentas de Google Cloud

1000

Horas hombre

2760 USD

Gastados en procesamiento

2 de 3 ensambles

en el podio

“Retreat...? ThunderCats **never** retreat!”

—Lion-O

Gracias!

Créditos

Feature Engineering	Fede Artur Maca
Organización General	Sofi Biam!
Scripts	Maca Sofi Biam!
Contención emocional	Thundercats
Mayordomo	El Kevin
Catering	El Viejo Palermo Dia%
Mascotas	Cucho Mina Gatita