

# ANFIBIOS

Andrea Carolina Menco Tovar  
Maestrante en estadística aplicada  
universidad Tecnológica de Bolívar



# DATOS ATÍPICOS

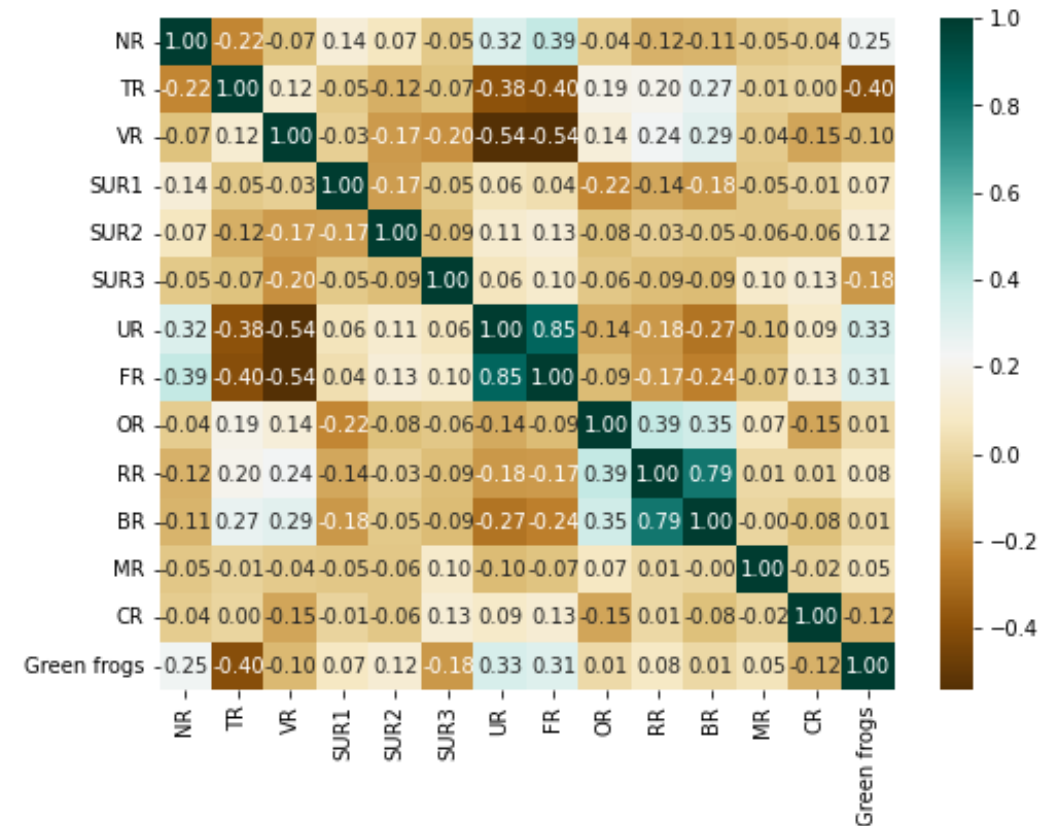
	NR	TR	VR	SUR1	SUR2	SUR3	UR	FR	OR	RR	BR	MR	CR	Green frogs
155	1	15	1	1	1	6	0	0	100	1	1	2	1	0

```
# Eliminar valores atípicos
data= data.drop(Outliers_to_drop, axis = 0).reset_index(drop=True)
data
```

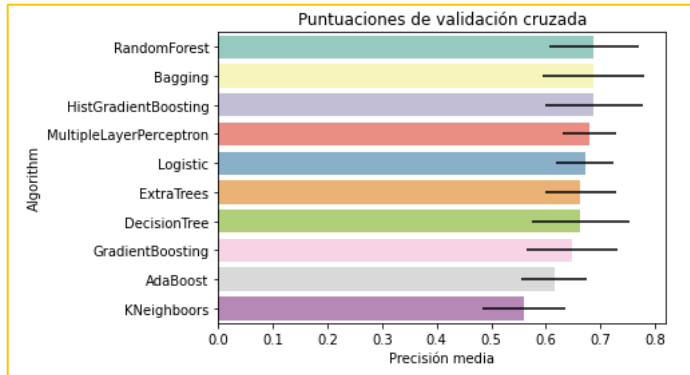
	NR	TR	VR	SUR1	SUR2	SUR3	UR	FR	OR	RR	BR	MR	CR	Green frogs
0	1	1	4	6	2	10	0	0	50	0	0	0	1	0
1	1	5	1	10	6	10	3	1	75	1	1	0	1	0
2	1	5	1	10	6	10	3	4	75	1	1	0	1	0
3	1	5	0	6	10	2	3	4	25	0	0	0	1	0
4	2	1	4	10	2	6	0	0	99	0	5	0	1	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
183	1	12	3	2	2	1	0	0	75	2	1	0	1	0
184	1	14	2	7	10	2	0	0	100	5	5	0	1	1
185	1	1	4	1	10	2	0	0	100	5	5	0	1	1
186	1	12	3	2	1	6	0	0	100	1	0	0	1	0
187	1	12	3	2	6	10	0	0	100	1	1	0	1	0

188 rows x 14 columns

# MATRIX DE CORRELACIÓN



# MODELO DE CLASIFICACIÓN

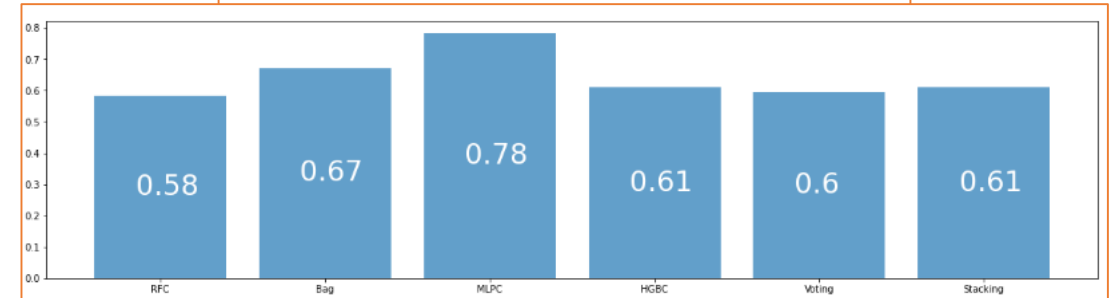


	CrossValMeans	CrossValerrors	Accuracy	Algorithm
0	0.696	0.074189	0.645161	RandomForest
5	0.688	0.064000	0.629032	HistGradientBoosting
4	0.672	0.099277	0.564516	Bagging
6	0.672	0.089084	0.693548	Logistic
2	0.664	0.078384	0.596774	GradientBoosting
3	0.648	0.053066	0.645161	ExtraTrees
8	0.648	0.064000	0.645161	DecisionTree
9	0.632	0.085417	0.709677	MultipleLayerPerceptron
1	0.624	0.082365	0.661290	AdaBoost
7	0.560	0.083905	0.612903	KNeighbors

RMSE for RFC is 0.5819876952473779  
 RMSE for Bag is 0.672021505032247  
 RMSE for HGBC is 0.6090712125322324  
 RMSE for MLPC is 0.7828813612588127

A partir de lo anterior observamos que el mejor RMSE es RandomForestClassifier debido a que presenta el más bajo 0.58.

	0	1	2	3	4	5
Modelo	RFC	Bag	MLPC	HGBC	Voting	Stacking
RMSE	0.581988	0.672022	0.782881	0.609071	0.595683	0.609071



RF 82  
 Bag 74  
 HGBC 85  
 MPL 42  
 Voting 84  
 Stracking 82

