



UNIVERSIDAD
NACIONAL
DE COLOMBIA

PROYECTO **CULTURAL, CIENTÍFICO Y COLECTIVO** DE NACIÓN

Variabilidad mensual de las velocidades de flujo en una zona del mar Caribe

Paula Andrea Espinosa Ordoñez

Universidad Nacional de Colombia

PROYECTO **CULTURAL, CIENTÍFICO Y COLECTIVO** DE NACIÓN

Datos

Variables Independientes (X):

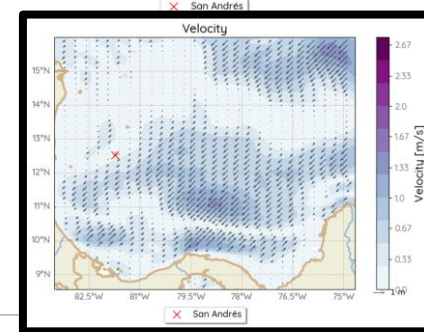
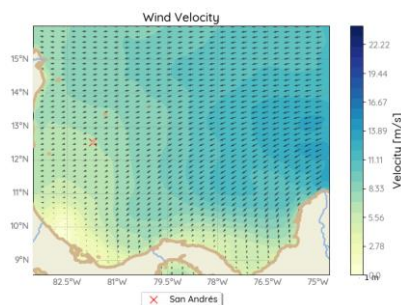
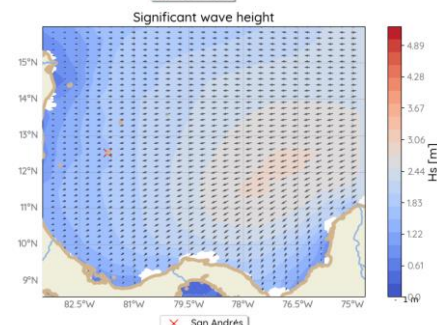
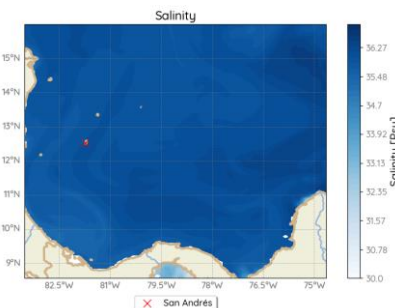
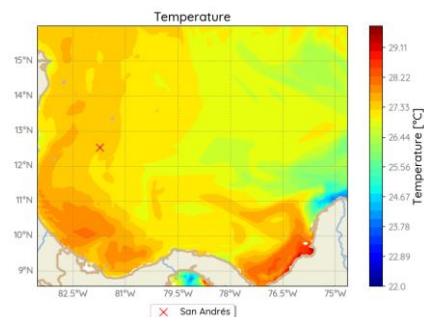
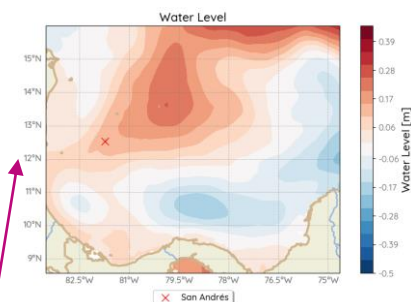
- Nivel del agua
- Salinidad
- Temperatura
- Altura y dirección de ola
- Magnitud y dirección del viento

Resolución:

- Píxeles: 4km x 8 km
- Temporal: promedio mensual
- Hycom
- ERA5

Variables Dependiente (Y):

- Magnitud de la velocidad
- Dirección de la velocidad



Configuración de la base de datos

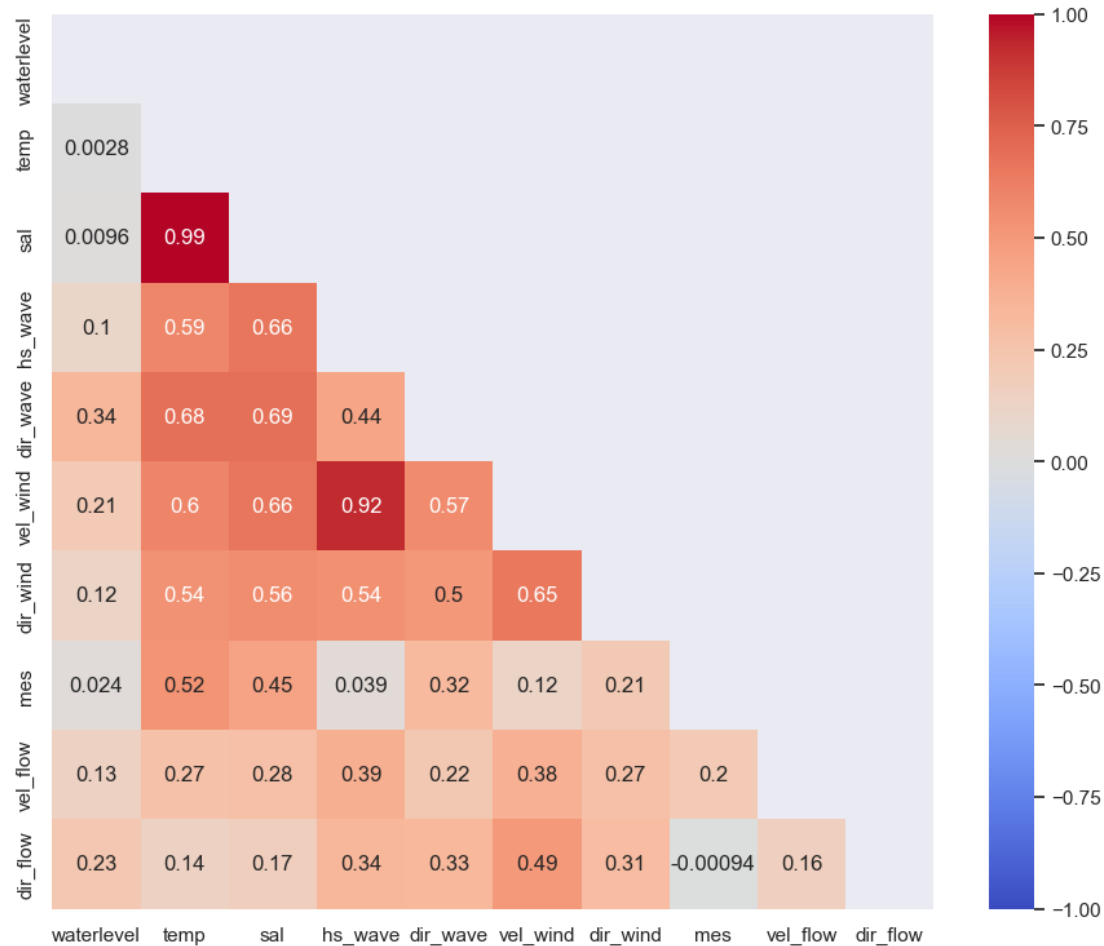
Información
del pixel

index	
0	46
1	47
2	48
3	49
4	50
...	...
216352	20939
216353	20940
216354	20941
216355	20942
216356	20943

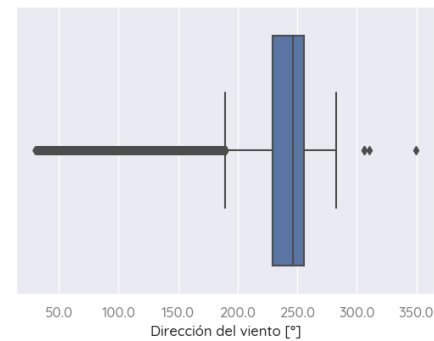
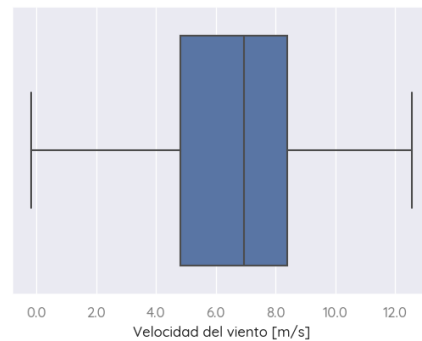
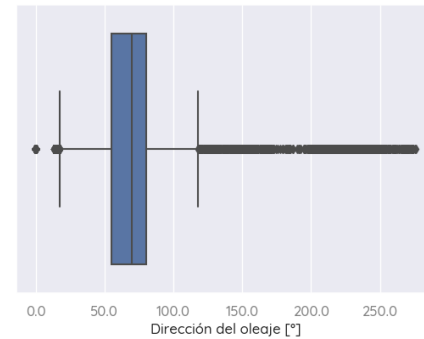
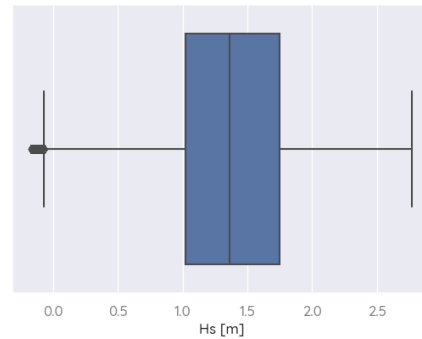
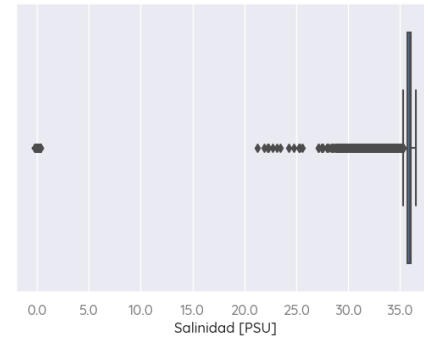
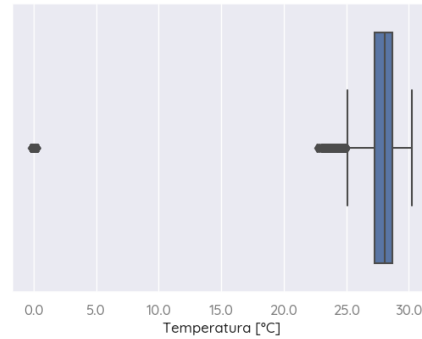
	waterlevel	temp	sal	hs_wave	dir_wave	vel_wind	dir_wind	mes	vel_flow	dir_flow
0	0.260220	0.260220	0.260220	0.116837	0.116837	0.116837	45.000000	1	0.368006	45.0
1	0.263163	0.263163	0.263163	0.122465	0.122465	0.122465	45.000000	1	0.372168	45.0
2	0.258976	0.258976	0.258976	0.130494	0.130494	0.130494	45.000000	1	0.366247	45.0
3	0.256825	0.256825	0.256825	0.140012	0.140012	0.140012	45.000000	1	0.363206	45.0
4	0.258829	0.258829	0.258829	0.149988	0.149988	0.149988	45.000000	1	0.366040	45.0
...
216352	0.218121	28.110162	35.755665	1.887047	87.019027	9.213161	259.490423	12	0.781339	225.0
216353	0.222498	28.105264	35.748895	1.888027	87.102830	9.229263	259.621510	12	0.802535	225.0
216354	0.226166	28.097134	35.737414	1.888999	87.187034	9.245683	259.752242	12	0.822391	225.0
216355	0.229069	28.081057	35.720507	1.889970	87.271393	9.257756	259.873328	12	0.838176	225.0
216356	0.231113	28.077571	35.710580	1.890939	87.355596	9.262286	259.979580	12	0.848334	225.0
216357 rows × 10 columns										



Análisis Exploratorio de datos



Análisis Exploratorio de datos



Selección de variables

Recursive Feature Elimination (RFE)

Magnitud



Nivel de agua
Temperatura
Salinidad
Altura de ola

Dirección

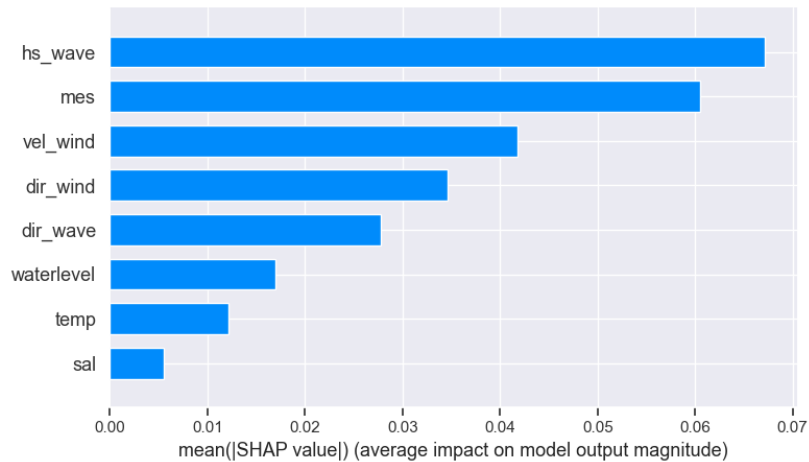


Nivel del agua
Temperatura
Altura de ola
Velocidad del viento

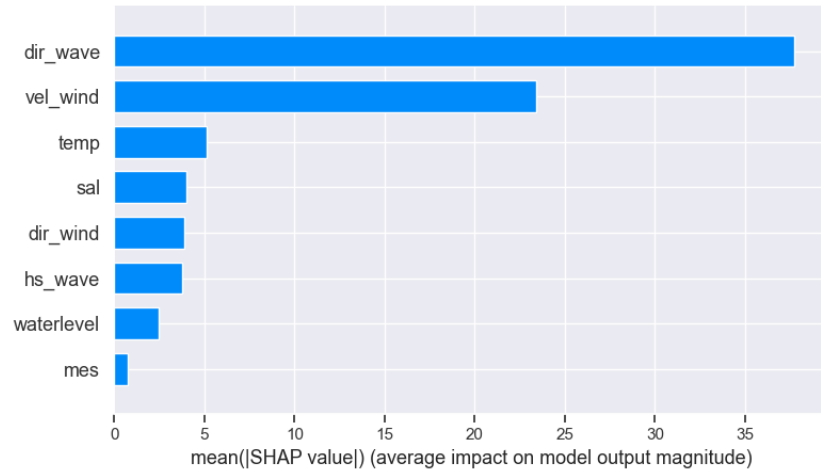


Selección de variables

Magnitud



Dirección



Validación Cruzada

K fold - Random Forest Regressor:

```
1 from sklearn.ensemble import RandomForestRegressor
2 #----- Para la magnitud-----
3 kfold = KFold(n_splits=5, shuffle= True, random_state=1)
4 model = RandomForestRegressor()
5 results = cross_val_score(model, X_train1, y_train1, cv=kfold, scoring='r2')
6 print(results)
7 print(results.mean())
8 print(results.std())
```

✓ 6m 35.1s

[0.96009712 0.95924487 0.95943987 0.96002153 0.95760486]
0.9592816500593436
0.0009000478611234187

```
1 #-----Para la dirección-----
2 kfold = KFold(n_splits=5, shuffle= True, random_state=1)
3 model = RandomForestRegressor()
4 results = cross_val_score(model, X_train2, y_train2, cv=kfold, scoring='r2')
5 print(results)
6 print(results.mean())
7 print(results.std())
```

✓ 6m 41.1s

[0.89903308 0.89348108 0.89896448 0.89675556 0.89421201]
0.896489240545718
0.002319638406659958

ShuffleSplit Linear Regression:

```
1 from sklearn.model_selection import ShuffleSplit
2
3 #----- Para magnitud-----
4 kfold = ShuffleSplit(n_splits=5, test_size=0.3, random_state=1)
5 model = LinearRegression()
6 results = cross_val_score(model, X1, Y1, cv=kfold, scoring='r2')
7 print(results)
8 print(results.mean())
9 print(results.std())
```

✓ 0.2s

[0.18924108 0.18736908 0.18806978 0.19148024 0.19229334]
0.1896907056713945
0.0019076513885836032

```
1
2 #----- Para la dirección-----
3 kfold = ShuffleSplit(n_splits=5, test_size=0.3, random_state=1)
4 model = LinearRegression()
5 results = cross_val_score(model, X2, Y2, cv=kfold, scoring='r2')
6 print(results)
7 print(results.mean())
8 print(results.std())
```

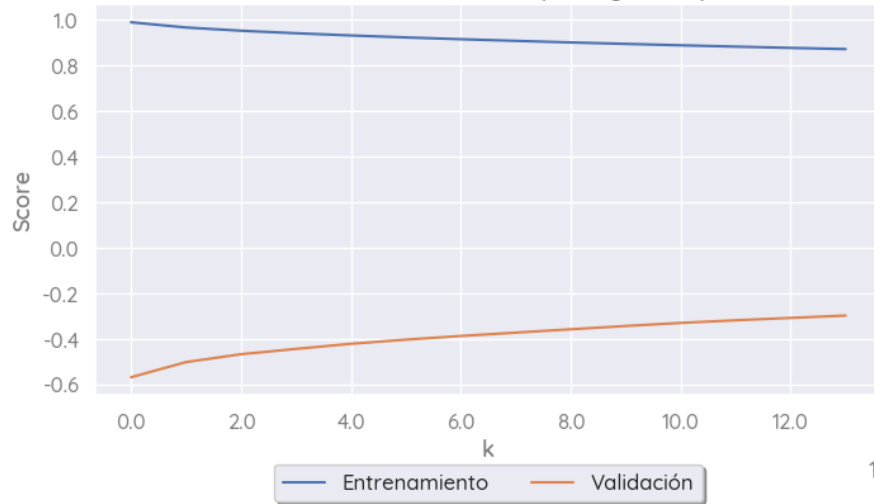
✓ 0.2s

[0.36135003 0.36334137 0.35528867 0.35955365 0.355551]
0.35901694116180266
0.003173161979158698

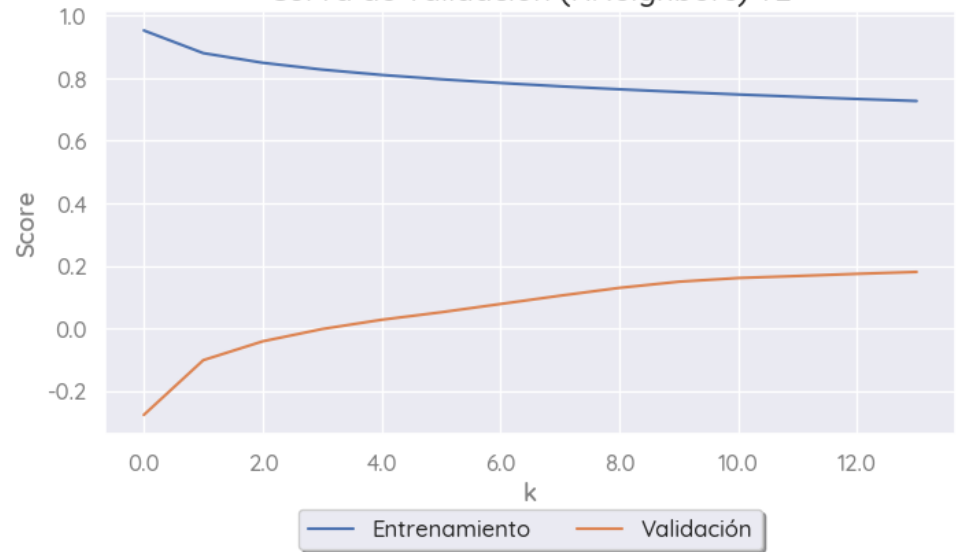


Curva de validación

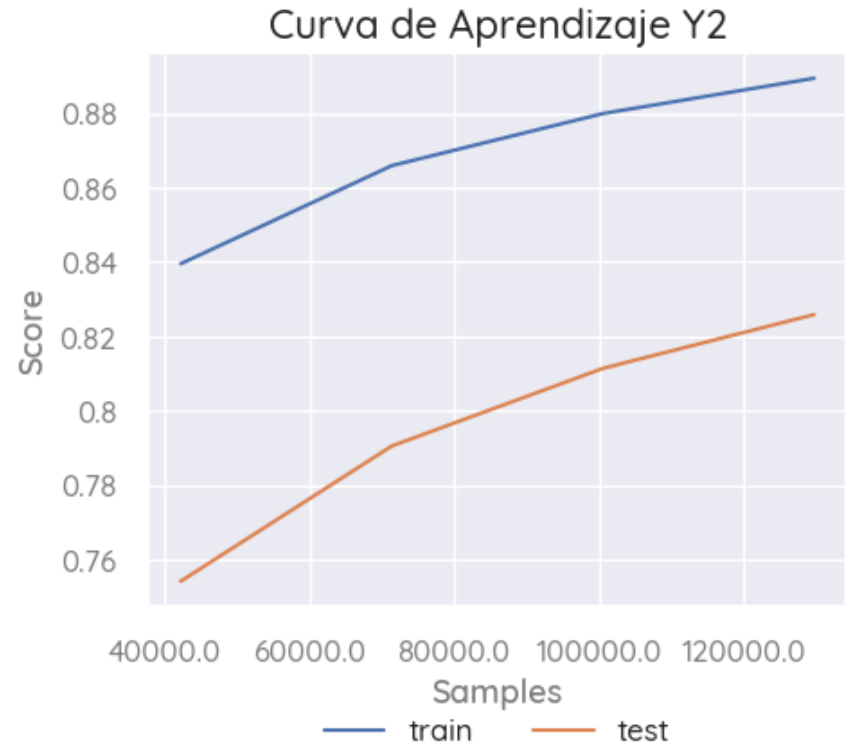
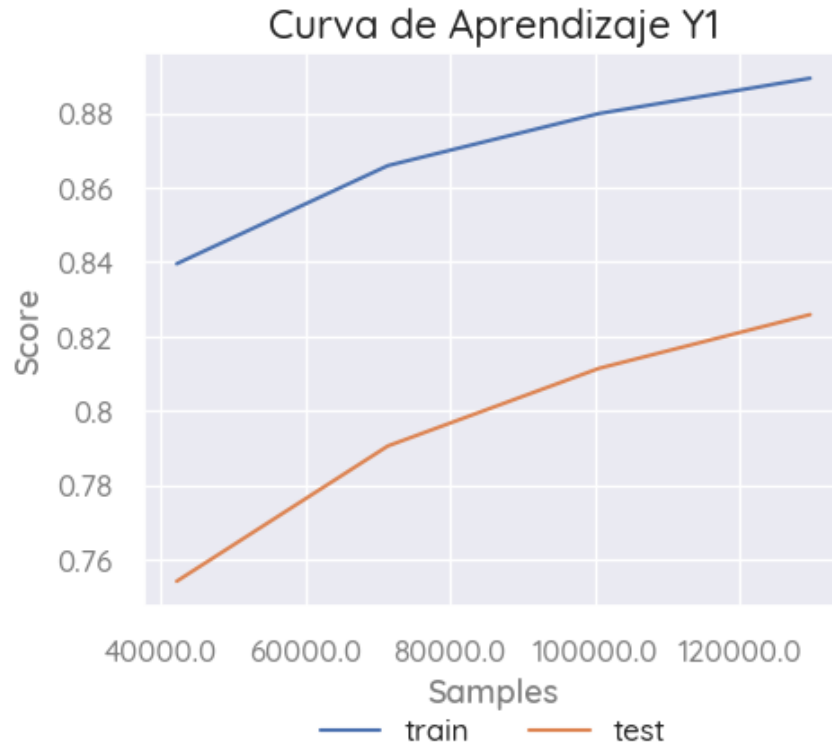
Curva de validación (KNeighbors)



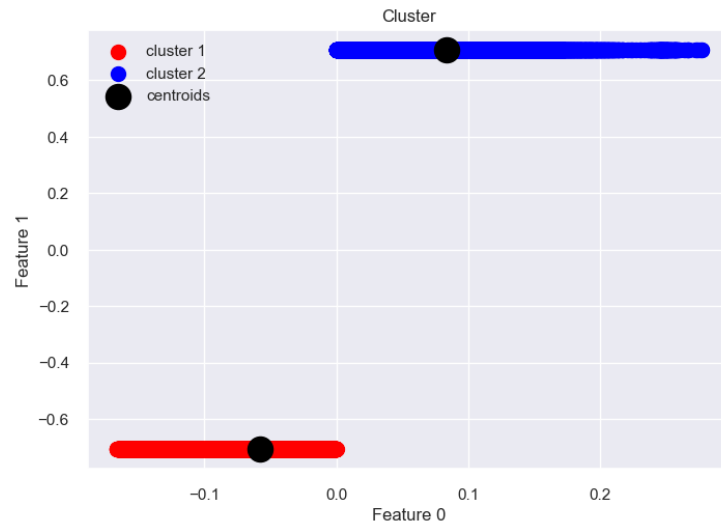
Curva de validación (KNeighbors) Y2



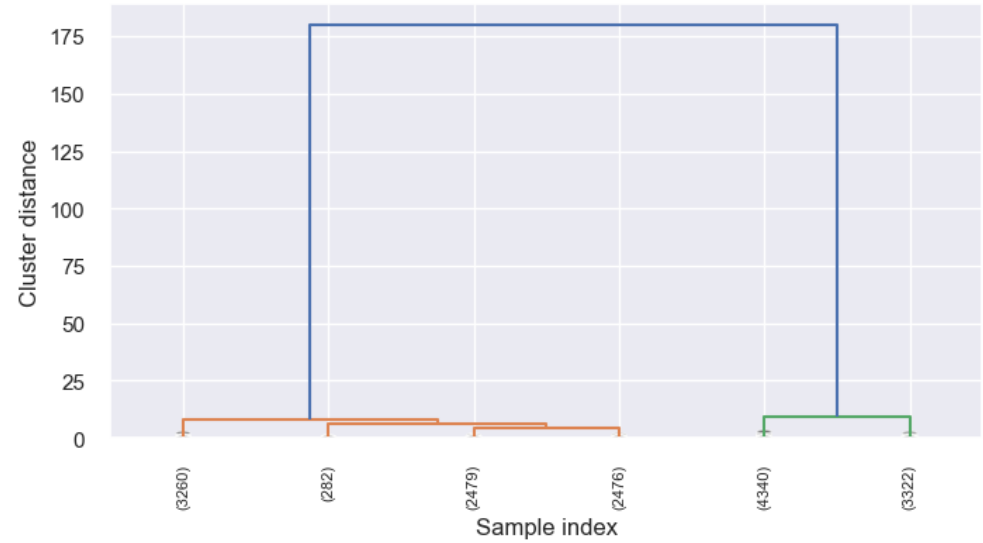
Curva de aprendizaje



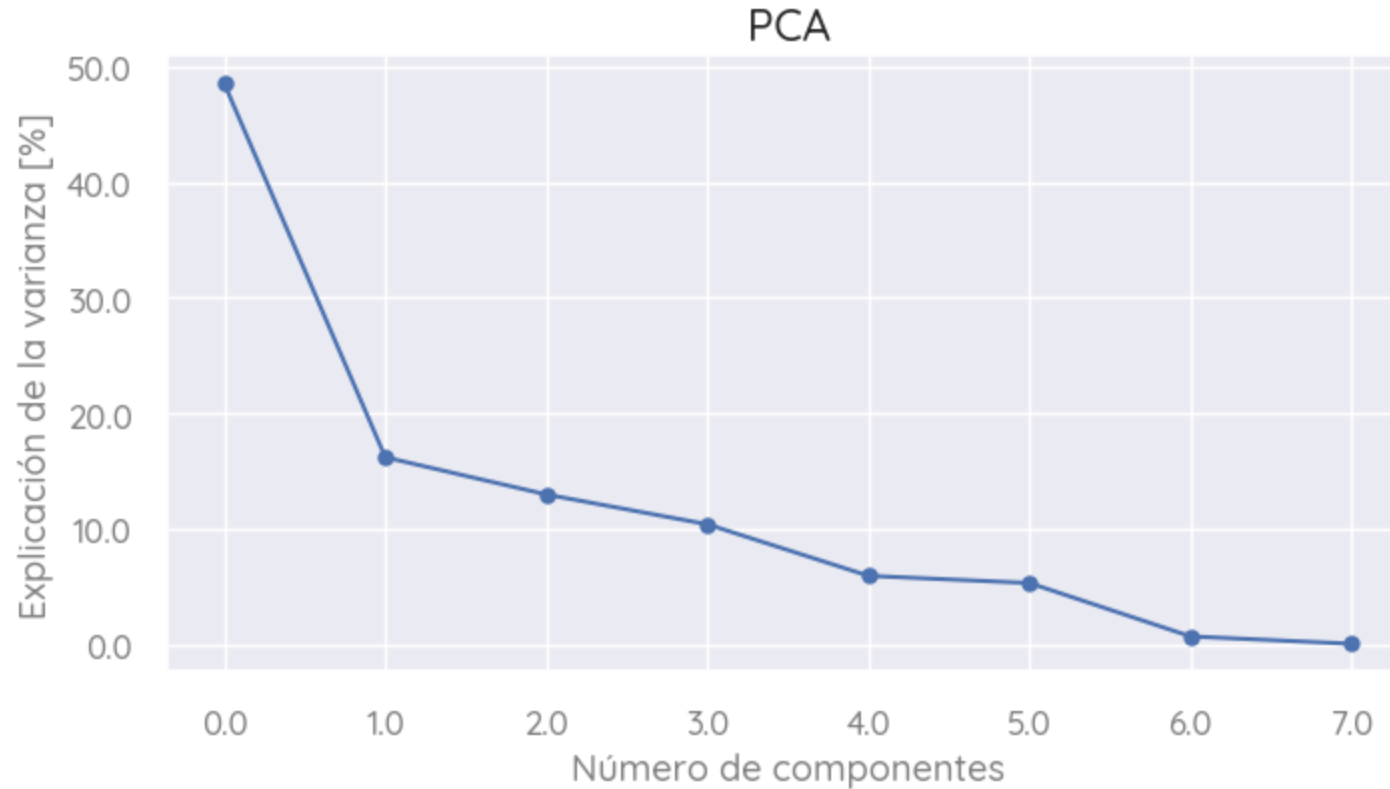
Clustering



Dendrograma



Análisis de componentes principales



Análisis de componentes principales



REFERENCIAS

- Cooper, K. M., & Barry, J. (2020). A new machine learning approach to seabed biotope classification. *Ocean & Coastal Management*, 198, 105361. <https://doi.org/10.1016/J.OCECOAMAN.2020.105361>
- de Clippele, L. H., Huvenne, V. A. I., Orejas, C., Lundälv, T., Fox, A., Hennige, S. J., & Roberts, J. M. (2018). The effect of local hydrodynamics on the spatial extent and morphology of cold-water coral habitats at Tisler Reef, Norway. *Coral Reefs*, 37(1), 253–266. <https://doi.org/10.1007/S00338-017-1653-Y/FIGURES/8>
- Lykkebo, Karem, Nadine Heck, Borja G. Reguero, Donald Potts, Armen Hovagimian, and Andina Paytan. 2019. “Biological and Physical Effects of Brine Discharge from the Carlsbad Desalination Plant and Implications for Future Desalination Plant Constructions.” *Water* 11:21.
- Olarte, Paloma Marina. 2019. “Climatología Del Transporte Potencial de Sedimentos Costeros Inducido Por Oleaje En La Isla San Andrés.” Universidad Nacional de Colombia.
- Escobar, Carlos A., Liliana Velásquez, and Federico Posada. 2015. “Marine Currents in the Gulf of Urabá, Colombian Caribbean Sea.” *Journal of Coastal Research* 31(6):1363–74.

AGRADECIMIENTOS

