

Elaborado por: Pasante de Ingeniería Bioquímica Vargas Carreño Leonardo Augusto

El Código completo del algoritmo "Receiver Operating characteristic" se encuentra descrito en el archivo ROC_BIOMECHANICAL.ipynb.

Resultados.

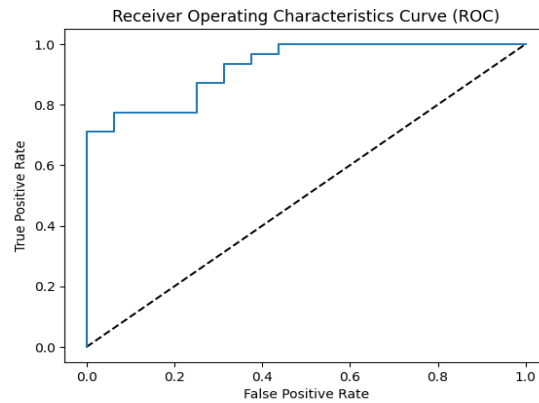
Tabla 1.1. Datos Generados ROC curve y AUC

Test Size	Cantidad Total de Datos		Cantidad de Datos Tomados				Tresholds Data	ROC AUC score	True Positive Rate	False Positive Rate
	X	Y	X train	Y train	X test	Y test				
0.15	310	310	263	263	47	47	Primeros Tres Datos inf 1. 0.8548569 0.85365055 0.83153443	0.925403	Primeros Tres Datos: 0.,0.0322580, 0.70967742 Últimos Tres Datos: 0.96774194, 0.96774194, 1., 1.	Primeros Tres Datos: 0., 0., 0. Últimos Tres Datos: 0.4375, 0.4375, 1.
0.50			155	155	155	155	Primeros Tres Datos: inf 1. 0.97844667 0.97825483 0.94085688	0.9269480	Primeros Tres Datos: 0.,0.0101010, 0.58585859 Últimos Tres Datos: 0.98989899,1, 1.	Primeros Tres Datos: 0., 0., 0. Últimos Tres Datos: 0.76785714, 0.76785714, 1.
0.95			15	15	295	295	Primeros Tres Datos: inf 1.00000000e+00 9.50152933e-01 9.49025568e-01 9.31194777e-01	0.8455232	Primeros Tres Datos: 0.,0.004950, 0.58910891 Últimos Tres Datos: 0.9950495, 0.9950495,1.	Primeros Tres Datos: 0., 0., 0. Últimos Tres Datos: 0.97849462, 1., 1.

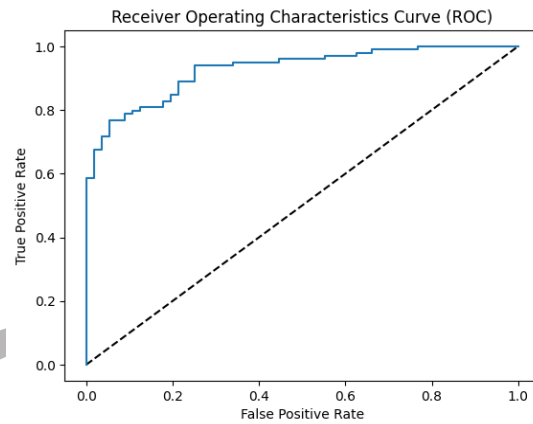
Elaborado por: Pasante de Ingeniería Bioquímica Vargas Carreño Leonardo Augusto

Gráficas True Positive Rate con respecto a False Positive Rate

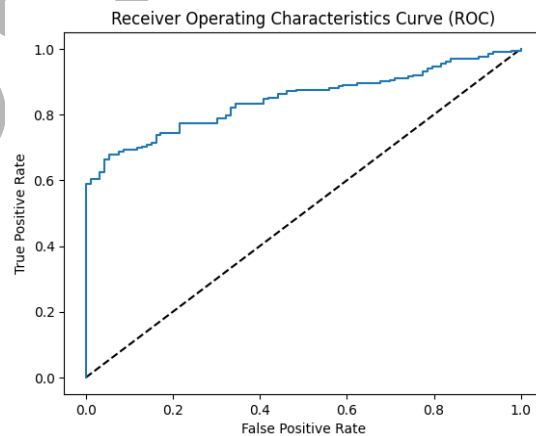
Test Size= 0.15



Test Size= 0.50



Test Size= 0.95



Elaborado por: Pasante de Ingeniería Bioquímica Vargas Carreño Leonardo Augusto

Discusión.

El algoritmo "ROC curve" es sencillo de implementar, es una simple curva que determina si el modelo de predicción reconoce las clasificaciones (principalmente se emplea para clasificaciones binarias, también, se puede emplear en clasificaciones con etiquetas múltiples). Para realizar la gráfica "ROC curve" se emplean los valores calculados TPR (True Positive Rate) y FPR (False Positive Rate). (Dipanja, Raghav, & Tushar, 2018) En el dataset Biomechanical tiene dos clases Abnormal (1) y Normal (0).

Se puede variar la cantidad de datos que se van a emplear en el entrenamiento y los datos de prueba para generar el algoritmo por lo que se pueden obtener diferentes resultados de la sensibilidad, como se puede observar Tabla 1.1. Datos Generados ROC curve y AUC con los diferentes valores de la variable Test Size. AUC Score varia desde 0.9254 hasta 0.8455 siendo valores aceptables para el algoritmo, si los valores de AUC fueran menores de 0.5 el modelo no tiene la capacidad de discriminación entre clase positiva y la clase negativa. Con AUC= 0, el modelo predice una clase negativa como positiva.

Conclusión.

Para el Dataset empleado el algoritmo para generar las gráficas ROC curve satisface las necesidades de predicciones para conocer si el modelo puede identificar las clases (Normal y Abnormal).

Bibliografía

Dipanja, S., Raghav, B., & Tushar, S. (2018). Chapter 5 Building, Tuning and Deploying Models. "Receiver Operating Characteristic". En *Practical Machine Learning with Python. A Problem-Solver's Guide to Building*. (pág. 276). Apress.