

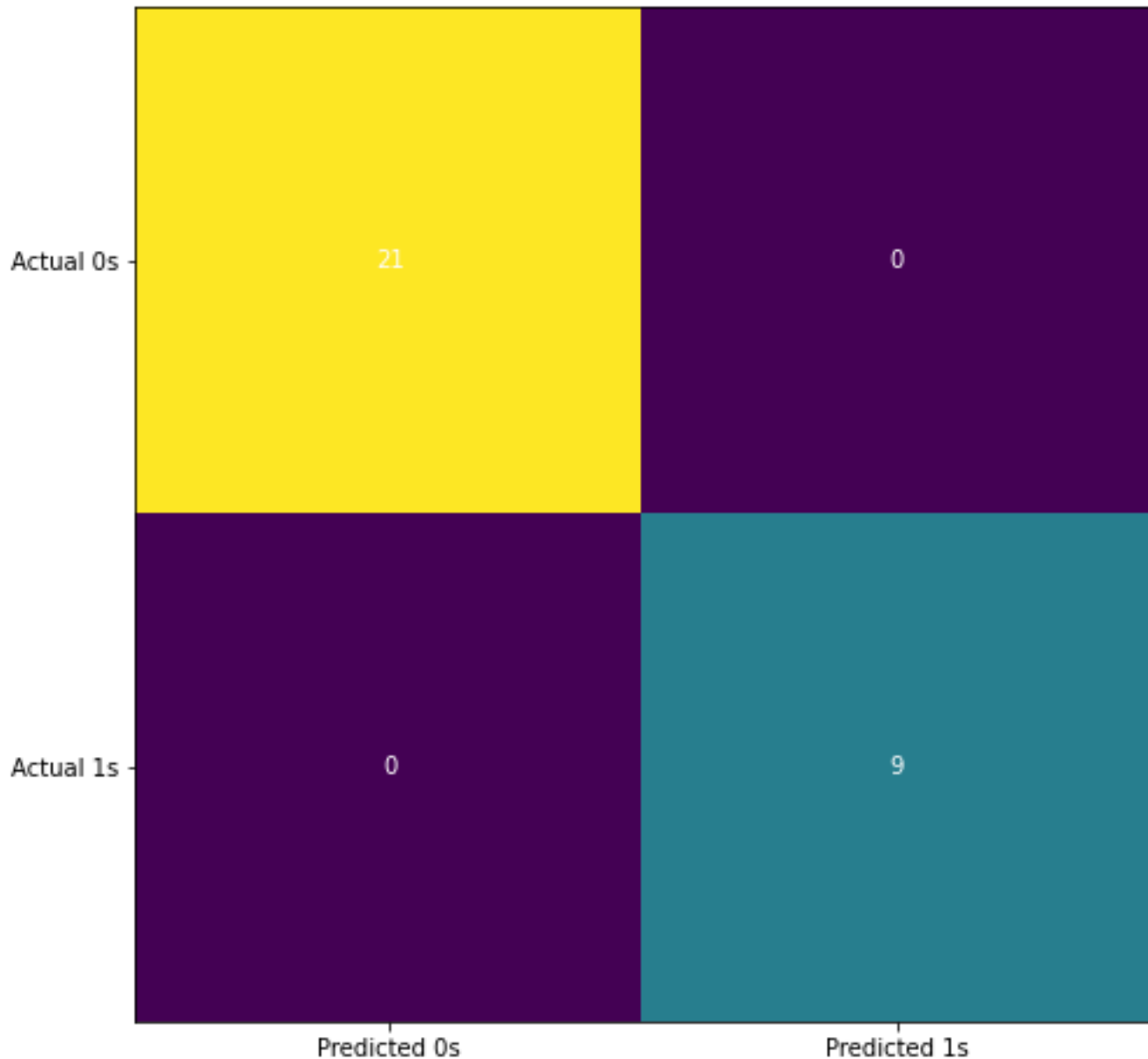
Alejandro Domínguez Lugo
A01378028

Momento de Retroalimentación

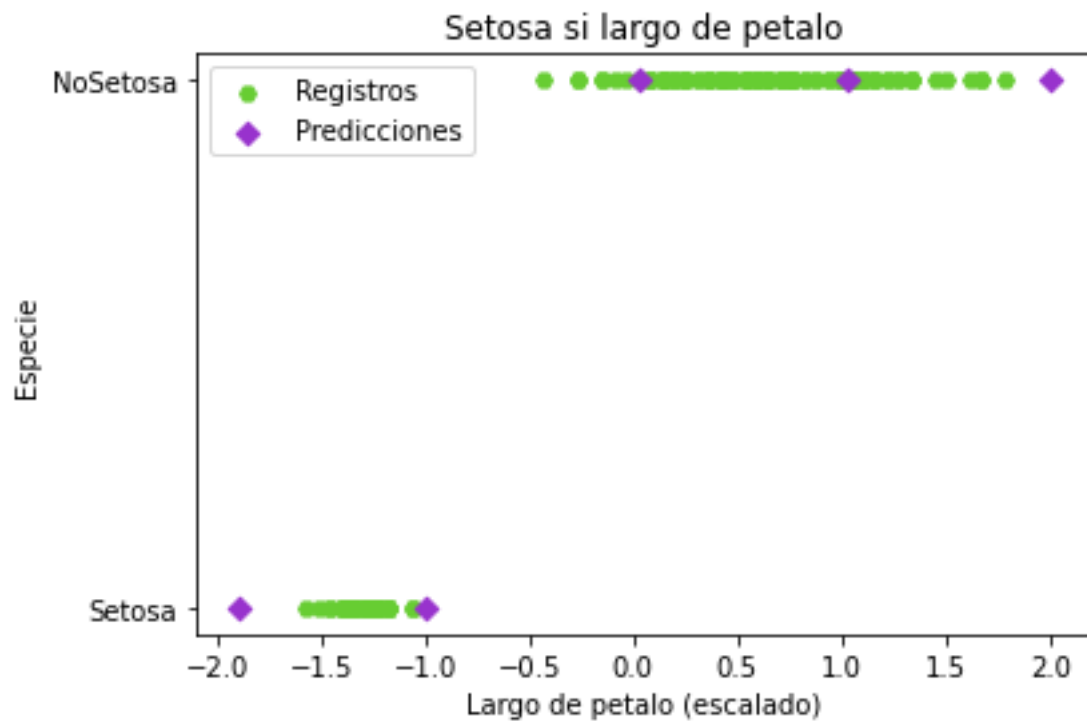
Módulo 2

Uso de framework o biblioteca de aprendizaje máquina para la implementación de una solución. Reporte

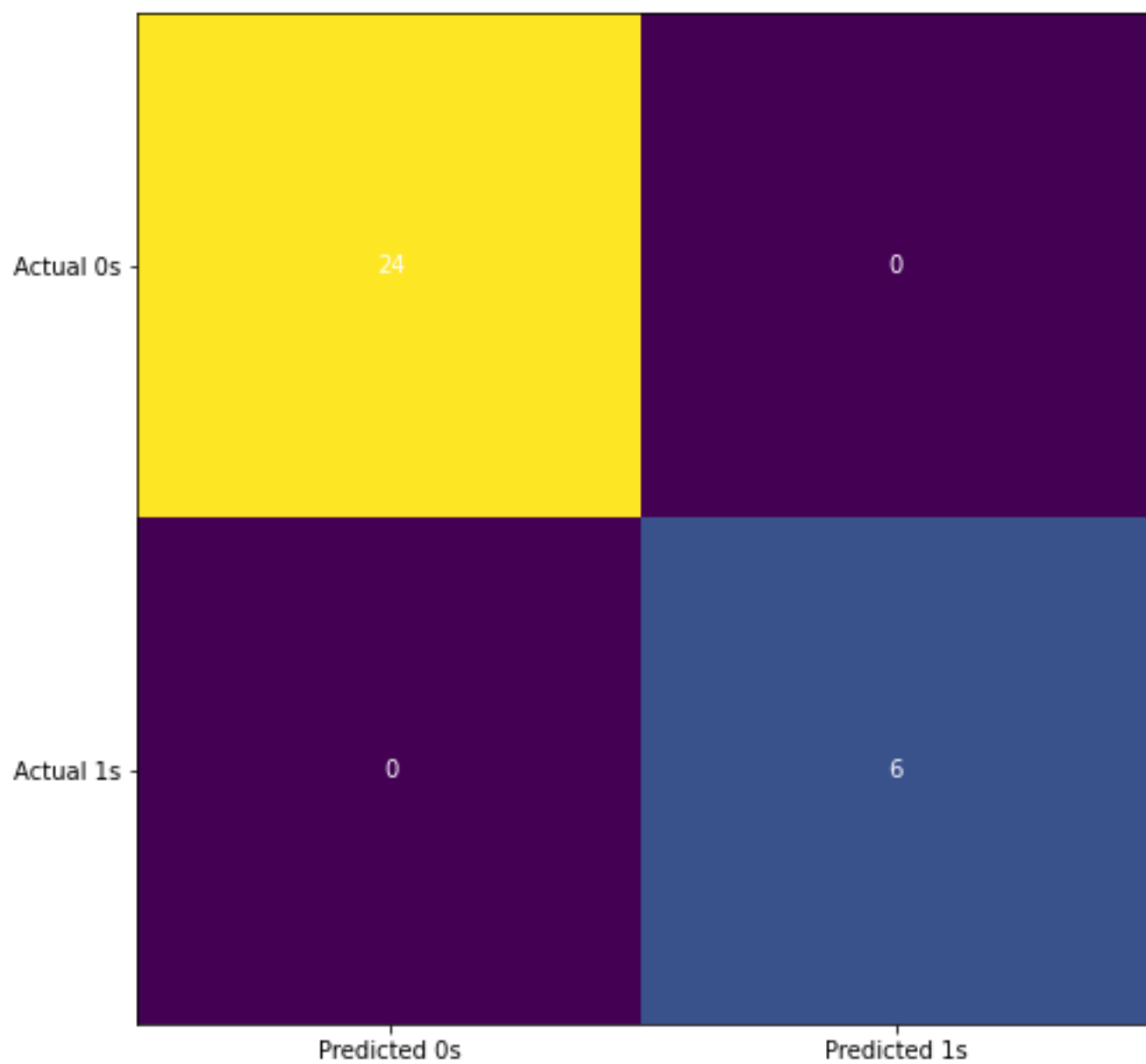
Con el código implementado se realizaron 9 pruebas para observar el comportamiento del modelo en distintas situaciones. En la primera prueba se probó una identificación sencilla de una flor si es setosa o no setosa con el largo del pétalo. Se observó que el modelo alcanzó la evaluación más alta tanto en entrenamiento como en evaluación y presentó la siguiente matriz de confusión.

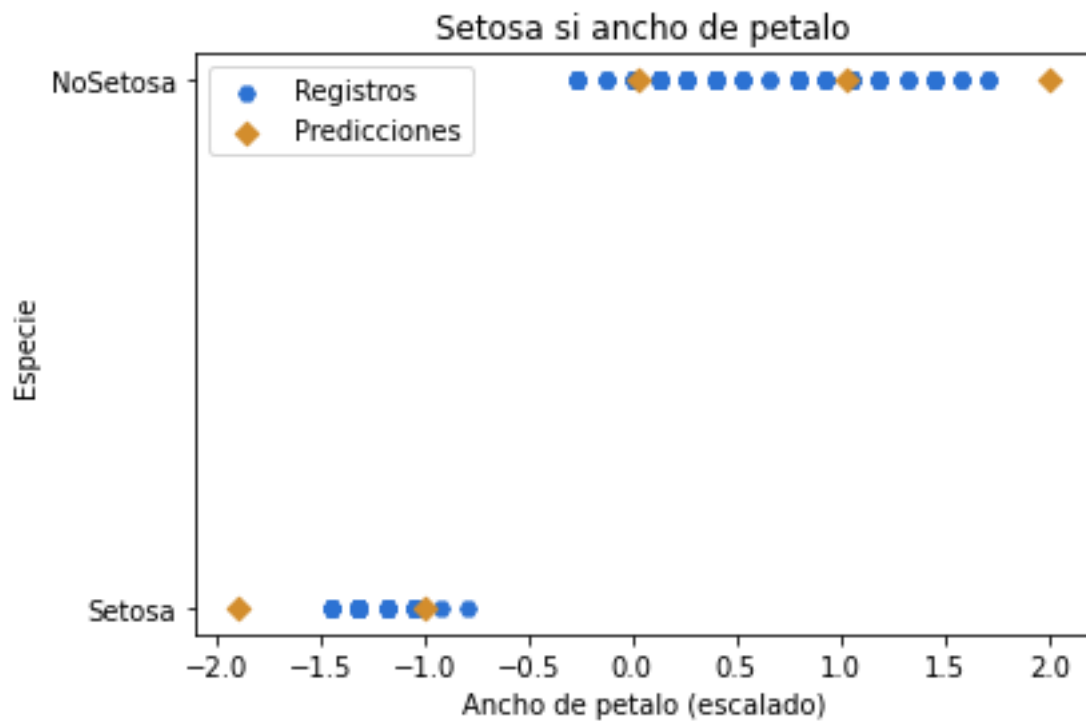


Posteriormente, se realizaron predicciones observando lo siguiente.

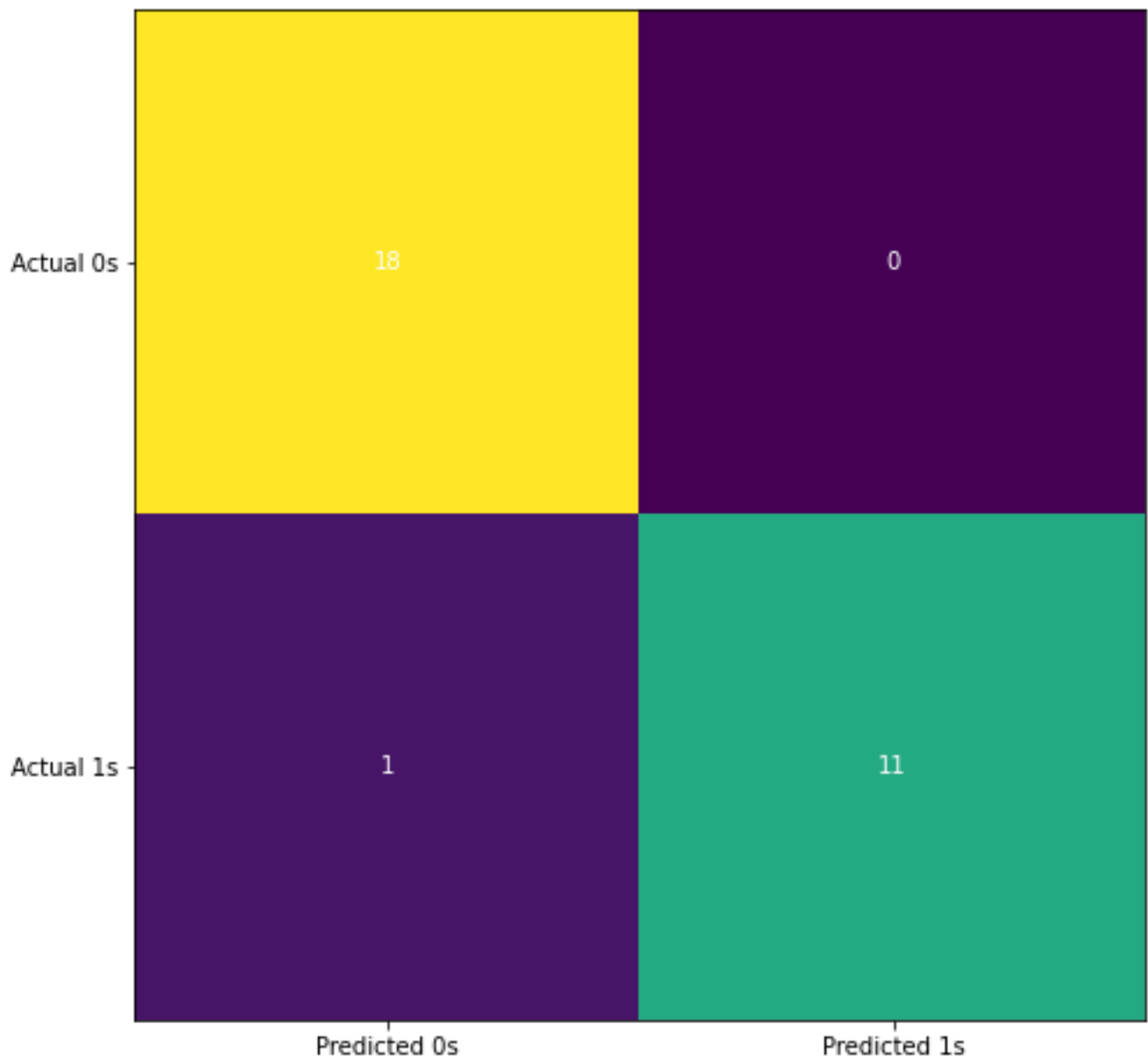


En la segunda prueba se utilizaron dos variables de entrada para realizar la identificación y el modelo resultó tener una evaluación excelente con la siguiente matriz de confusión y predicciones.

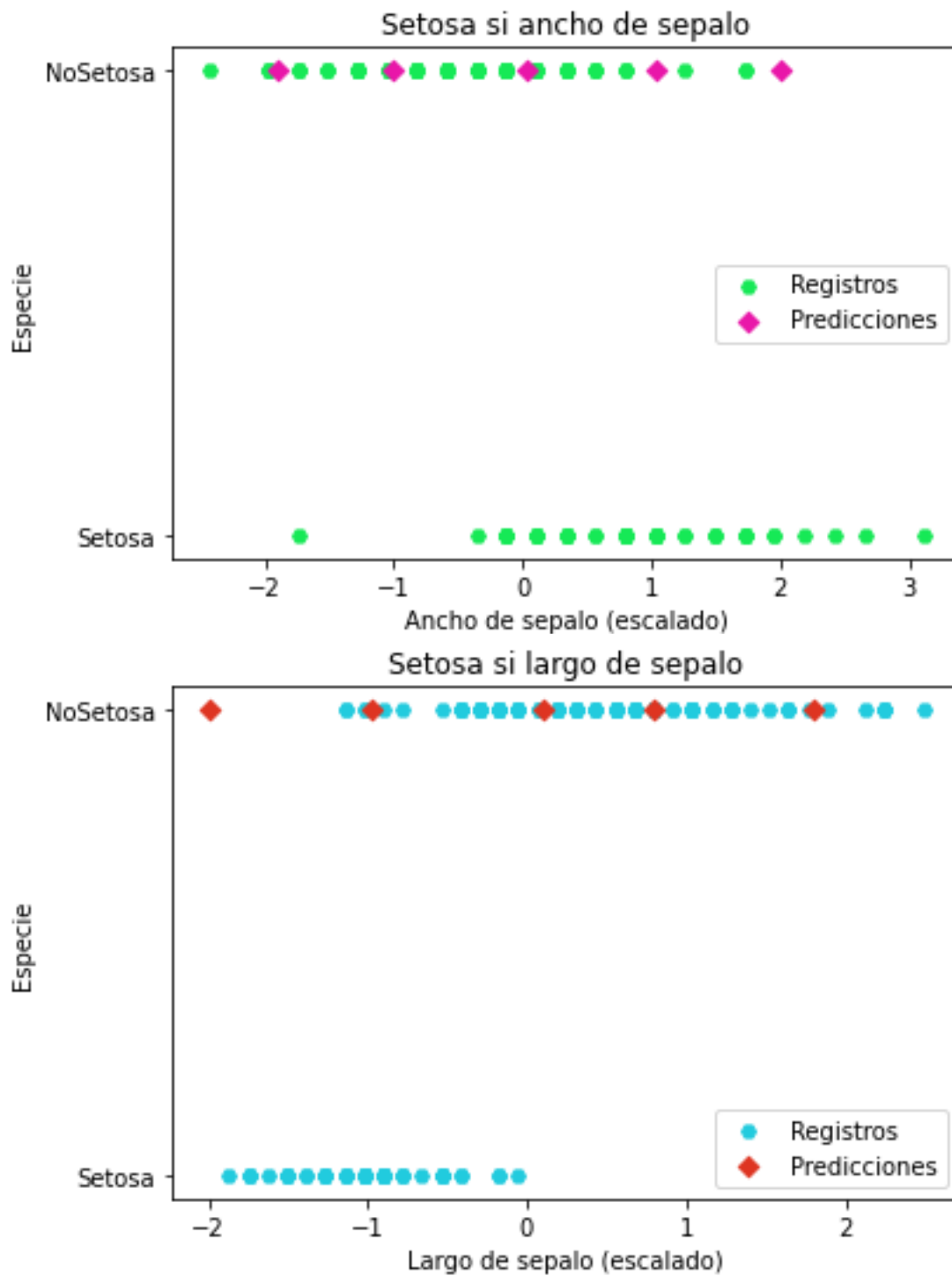




En la tercera prueba se utilizaron las otras dos variables de entrada y estas lograron causar errores en el modelo que alcanzó un 1 de calificación en entrenamiento, pero un 0.97 en validación. En la matriz de confusión se puede observar que identificó una flor como no setosa cuando lo era.

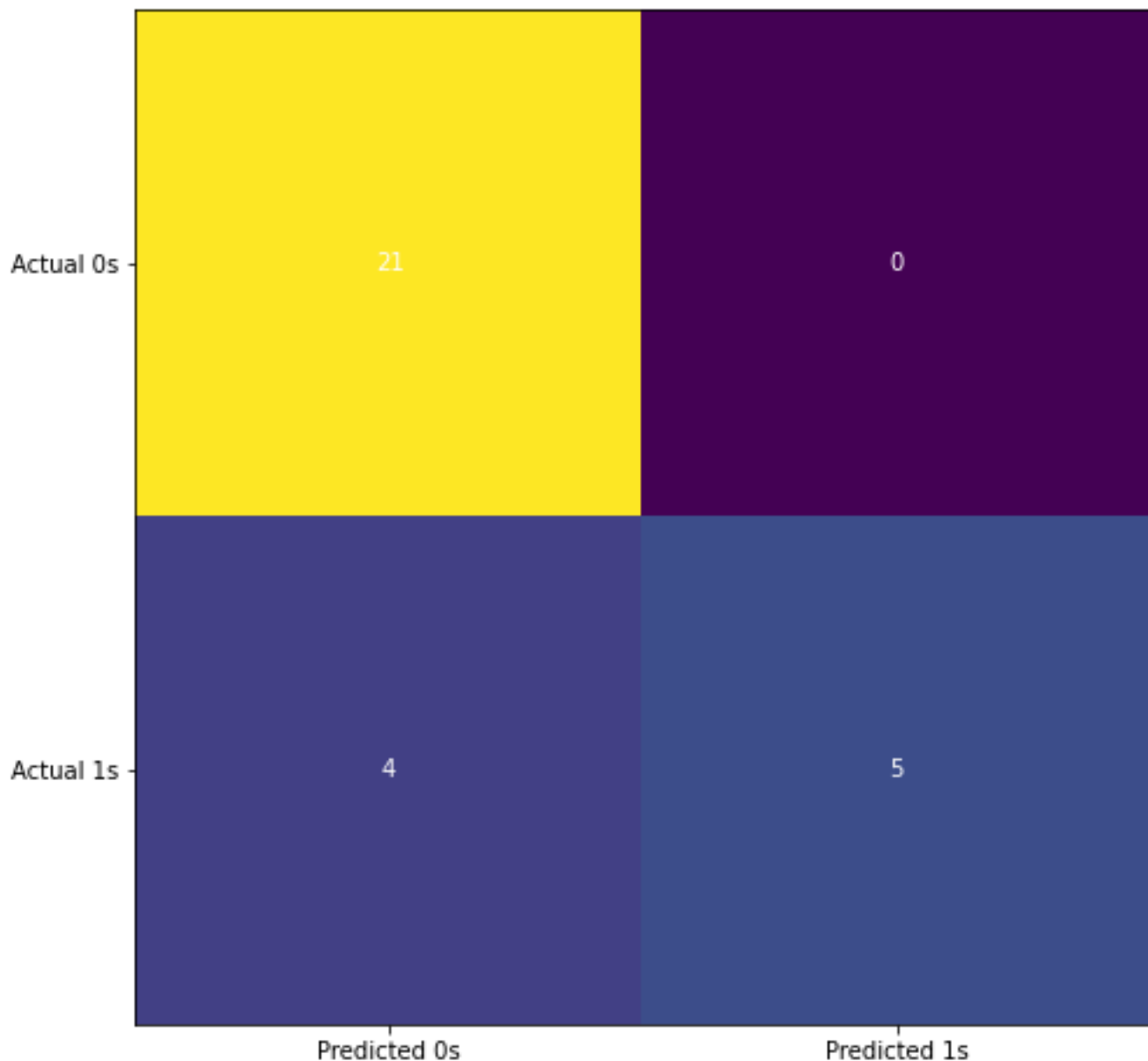


Con las predicciones se puede observar por qué el modelo no es perfecto.



Resulta ser que los registros se sobreponen.

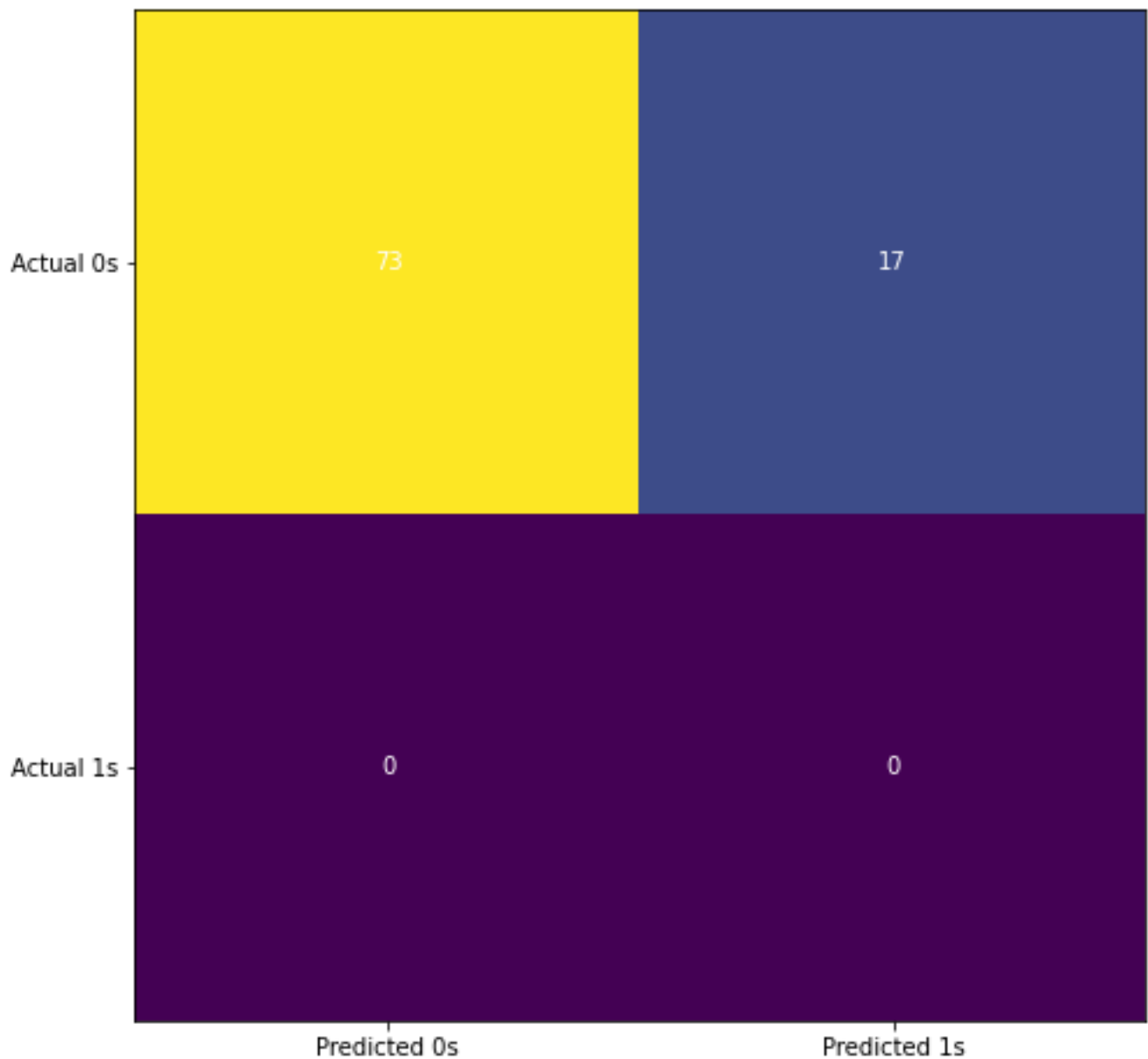
En la cuarta prueba se buscó enfocarse en esta sobreposición utilizando solo el ancho del sepal. El modelo obtuvo un puntaje de 0.82 en entrenamiento y de 0.87 en validación, con la siguiente matriz de confusión.



En la quinta prueba se utilizaron todas las variables de entrada y el resultado fue un modelo con evaluación de 1 en entrenamiento y validación.

En la sexta prueba se utilizaron las entradas relacionadas con el sepalo y se variaron los hiperparámetros alcanzando un modelo con evaluación perfecta. En la séptima prueba se utilizaron las mismas entradas, pero con distintos hiperparámetros alcanzando otro modelo con evaluación perfecta. Para la octava prueba, por el otro lado, se buscó generar un modelo poco preciso modificando los hiperparámetros para demostrar que, si son modificados sin saber, los resultados pueden no ser los deseados.

Este ultimo modelo obtuvo una calificación en entrenamiento de 0.95 y de 0.81 en validación con la siguiente matriz de confusión.



Por último, en la prueba final se dejaron la salida con las tres especies distintas de flor para verificar que la regresión logística podría lograr una clasificación no binaria. Se obtuvo un modelo con 0.97 de calificación en entrenamiento y de 0.93 en validación. Se realizaron predicciones y este fue el resultado.

