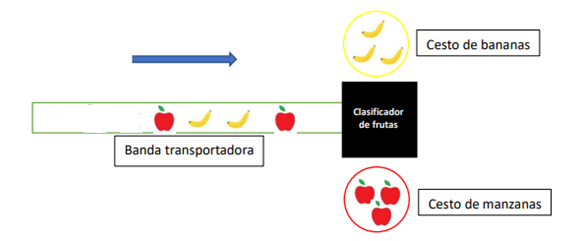
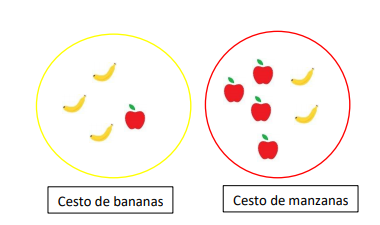
***Taller de evaluación de modelos de Machine Learning***

Hay un modelo de machine Learning que clasifica entre manzanas y bananas el cual fue implementado dentro de un sistema de bandas transportadoras como se muestra en la siguiente gráfica:

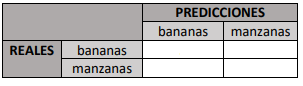


Las diez frutas fueron clasificadas de la siguiente manera:



Se pide:

1. Calcular la matriz de confusión.



|  |  |  |  |
| --- | --- | --- | --- |
|  | | PREDICCIONES | |
| BANANAS | MANZANAS |
| REALES | BANANAS | 3 | 2 |
| MANZANAS | 1 | 4 |

1. Calcular Tasa de las predicciones hechas correctamente respecto a todos los valores evaluados. (Todas las frutas).

Accuracy =

Accuracy = 3+4/3+4+1+2 = 0,7

1. ¿Qué proporción de frutas detectadas como Bananas en realidad si fueron Bananas?

Precision = VP / (VP + FP)

Precision = 3/3+1 = 0,75

1. ¿Qué proporción de frutas que realmente son Bananas fueron detectadas como Banana?

Recall = VP/ (VP + FN)

Recall = 3/3+2 = 0,6

1. Calcula una métrica que tenga en cuenta la precisión y la sensibilidad del modelo.

F1-Score = (2 \* Precision\*Recall) / (Precision + Recall)

F1-Score = 2\*0,75\*0,6/0,75+0,6 = 0,66

1. Supongamos que usted tiene dos modelos que reconocen si una transacción tiene o no tiene fraude (Los modelos se llaman *clasificador A* y *clasificador B)*. Dichos modelos fueron entrenados con muchas transacciones etiquetadas como “Sin fraude” y “Con fraude”. Las métricas que se tienen son:

|  |  |  |
| --- | --- | --- |
| ***Clasificador*** | ***Precision*** | ***Recall*** |
| ***A*** | 95% | 90% |
| ***B*** | 98% | 85% |
|  |  |  |

En este caso, para el *clasificador A*, hay un 95% de probabilidad de que haya un fraude en una transacción y un 90% de probabilidad de que haya detectado correctamente un fraude. Mientras que para el clasificador B hay un 98% de probabilidad de que haya un fraude en la transacción y un 85% de probabilidad de que haya detectado correctamente un fraude.

¿Qué modelo elegiría usted?, ¿Cuál sería su recomendación?

Yo recomendaría el modelo A porque comparando con el B hay menos probabilidad de que haya un fraude y más probabilidad de en caso de que haya uno detectarlo.

F1-Score-A = 2\*0,95\*0,90/0,95+0,90 = 0,92

F1-Score-B = 2\*0,98\*0,85/0,98+0,85 = 0,91

También al ver el F1-Score el de A es mejor que el de B.

1. ¿Qué puede estar pasando en un modelo que tiene un 99% de precisión pero que NO detecte el 95% de los casos?

Pasa que le modelo aprendió de memoria en base a los datos con los que lo entrene, por lo que, al enfrentarlo a la vida real, el modelo por su sobreajuste no es capaz de analizar dichos datos nuevos y puede caer en esos problemas.