



**Instituto Tecnológico y de Estudios  
Superiores de Monterrey**  
Campus Puebla

**Inteligencia Artificial para la ciencia de datos TC3007C**

**Reto**  
**Momento de Retroalimentación: Reto Evaluación**

Fernando Jiménez Pereyra	A01734609
Daniel Flores Rodríguez	A01734184
Alejandro López Hernández	A01733984
Daniel Munive Meneses	A01734205

14 de noviembre de 2022

## Métricas a considerar

### *Accuracy*

La fracción de predicciones que el modelo realizó correctamente. Se representa como un porcentaje o un valor entre 0 y 1. Es una buena métrica cuando tenemos un conjunto de datos balanceado, esto es, cuando el número de etiquetas de cada clase es similar.

### *Loss*

Una función de pérdida, o Loss function, es una función que evalúa la desviación entre las predicciones realizadas por la red neuronal y los valores reales de las observaciones realizadas durante el aprendizaje.

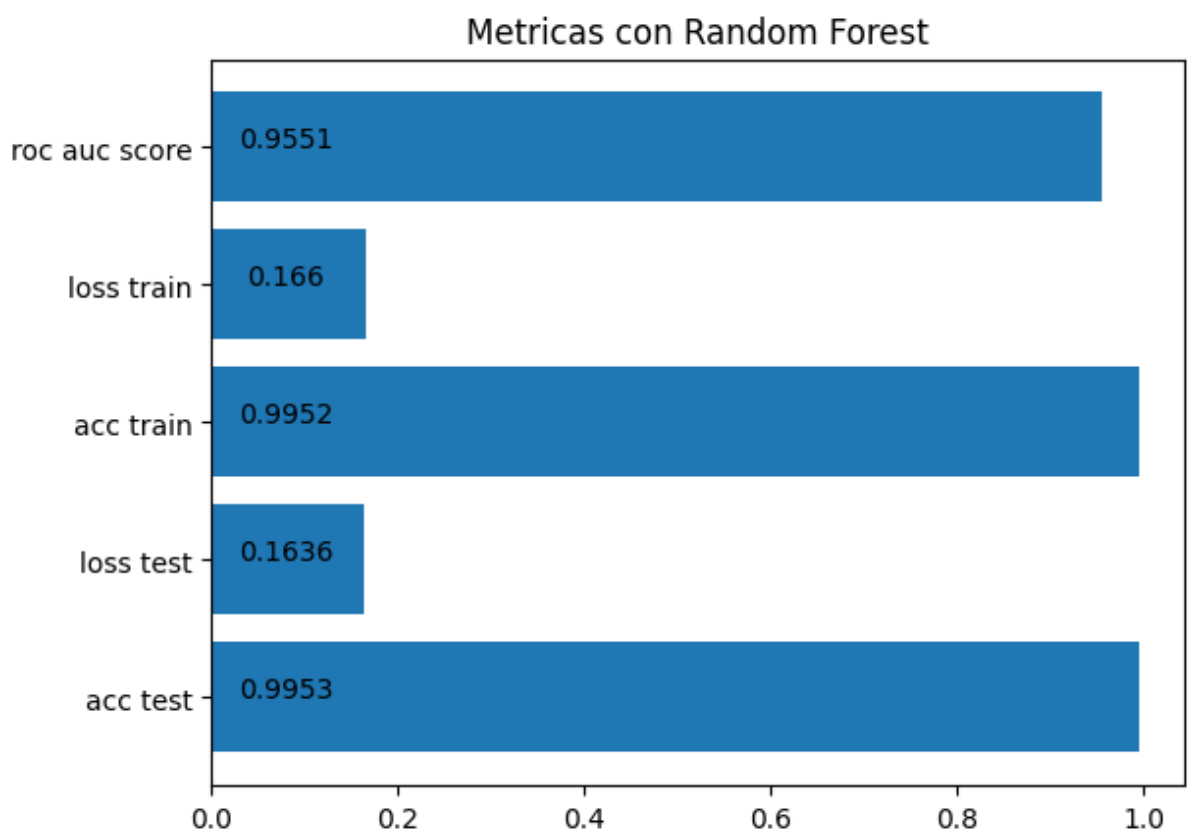
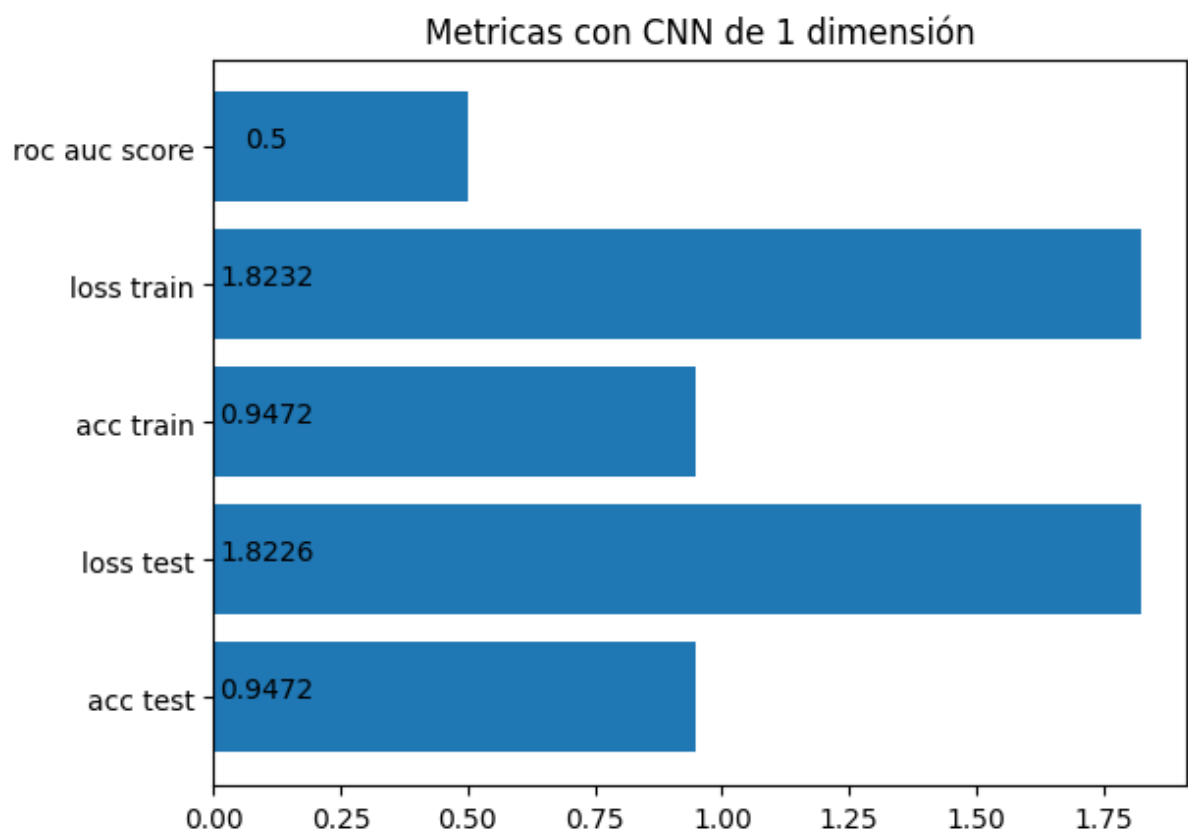
### *ROC-AUC Curve*

Una curva ROC es una representación gráfica de la sensibilidad frente a la especificidad para un modelo de clasificación. Es la representación de la razón o proporción de verdaderos positivos respecto a la razón o proporción de falsos positivos.

## Tabla de comparación de modelos respecto a las métricas seleccionadas

Modelo	Accuracy test	Accuracy train	Loss test	Loss train	ROC-AUC Curve
Random Forest	0.9952	0.9951	0.1636	0.1636	0.9551
CNN 1D	0.9472	0.9472	1.8226	1.8232	1

## Graficación



## Conclusión

Si bien el CNN presenta resultados con unos coeficientes casi perfectos, el Random Forest parece tener un comportamiento más realista y eficiente sin la necesidad de aplicar balanceo de datos. En adición a ello, se solicita que el sistema planteado para el reto contemple reajustar el modelo sobre la plataforma en caso de recibir set de datos diferentes; por lo que el rendimiento para el ajuste de los modelos en términos de tiempo resulta fundamental. El CNN, con el estado en el que está, tarda poco más de media hora en realizar el ajuste; mientras que el RandomForest tarda apenas unos minutos. En resumen, por términos de practicidad se opta por emplear el Random Forest sobre una arquitectura de deep learning como lo puede ser una red convolutiva de 1 dimensión.