

concordance=TRUE

## Resumen

En este trabajo se aborda la construcción de modelos de clasificación multiclase usando el conjunto de datos MNIST que es una gran base de datos de dígitos escritos a mano. Las metodologías que se abordarán para construir los modelos son las siguientes

1. Regresión logística multinomial
2. Árboles de clasificación
3. Bosques aleatorios
4. Máquinas de soporte vectorial
5. Redes neuronales

En este desarrollo se comparará el rendimiento de los modelos antes mencionados para verificar cuál es la mejor de estas metodologías con este conjunto de datos particular.

## 1. Definiciones

### MNIST

La base de datos MNIST (base de datos modificada del Instituto Nacional de Estándares y Tecnología) es una gran base de datos de dígitos escritos a mano que se usa comúnmente para capacitar a varios sistemas de procesamiento de imágenes. La base de datos también se usa ampliamente para capacitación y pruebas en el campo del aprendizaje automático.

### Modelo Estadístico

El modelado estadístico es una forma simplificada, matemáticamente formalizada, de aproximarse a la realidad, la que genera los datos) y, opcionalmente, hacer predicciones a partir de dicha aproximación.

## 2. Metodología

### 2.1. Análisis descriptivo

```
trainingData <- read.csv("../data/mnist_train.csv",header = FALSE)
table(trainingData$V1)
```

```
##
##      0      1      2      3      4      5      6      7      8      9
## 5923 6742 5958 6131 5842 5421 5918 6265 5851 5949
```

**2.2. Modelamiento estadístico**

**2.3. Diagnóstico del Modelo**

**3. Conclusiones**

**4. Referencias**