



Universidad Nacional de Rosario  
Facultad de Ciencias Exactas, Ingeniería y Agrimensura  
T.U.I.A  
Aprendizaje Automático II

## Trabajo Práctico 2

Redes Recurrentes  
Reinforcement Learning

1er cuatrimestre 2025

Autor/es:

Grupo N°	
Nombre y Apellido	N° de Legajo

Corrigió	Calificación

# Problema 1 - Audio MNIST

## Descripción:

En este problema, se presenta un conjunto de datos que contiene clips de audio correspondientes a dígitos hablados del 0 al 9.

## Dataset:

[https://www.tensorflow.org/datasets/catalog/spoken\\_digit](https://www.tensorflow.org/datasets/catalog/spoken_digit)

El dataset proporcionado incluye un total de 2500 clips de audio correspondientes a 5 locutores distintos, 50 clips por dígito por locutor.

## Objetivo:

Utilizando el dataset proporcionado, el objetivo es construir un modelo de clasificación utilizando redes neuronales que pueda inferir con precisión el dígito correspondiente dado un clip de audio. Se deben entrenar y evaluar modelos utilizando técnicas adecuadas de validación y métricas de evaluación de clasificación.

Se solicita entrenar dos modelos de distintas arquitecturas y comparar los resultados:

- Modelo convolucional sobre los espectrogramas de los clips.
- Modelo recurrente sobre los espectrogramas de los clips.

Ver <https://colab.research.google.com/github/FCEIA-AAII/lab11/blob/master/lab11-a.ipynb> como ejemplo de obtención de espectrogramas a partir de clips de audio.

## Entrega:

La entrega debe incluir:

Código fuente de la solución implementada en Google Colab, que incluya:

- Análisis previo y preprocesamiento del set de datos.
- Definición y entrenamiento del modelo.
- Resultados de la evaluación de los modelos, incluyendo métricas de desempeño y visualizaciones relevantes.

Nota: el código debe estar debidamente documentado con comentarios explicativos para que el trabajo sea fácilmente comprensible para otros revisores.