

Trabajo Práctico 01: Redes Neuronales

Objetivo: Crear y entrenar dos redes de tipo Perceptrón Multicapa con *BackPropagation* utilizando la biblioteca de Python [TensorFlow](#). Los datos para entrenar la red son provistos por la cátedra y se especifican más abajo.

Forma de entrega: El trabajo se deberá entregar por e-mail y deberá incluir:

- El archivo “.py” que corresponde a las redes neuronales, escrita en código python. O bien el enlace a la notebook de colab.research.google.com
- Un informe breve que indique:
 - Arquitectura de cada perceptrón creada para cada problema.
 - Qué casos se usaron para probar y cuales se usaron para testear (usar el 33% del conjunto de entrenamiento para validar la red)
 - Resultados obtenidos (en el conjunto de pruebas):
 - % de casos de éxito en el caso de la red de clasificación.
 - Error cuadrático medio para el caso de regresión
 - Iteraciones en cada caso

Fecha de entrega: deberá entregarse a más tardar el día 8 de mayo.

Datos:

Los datos podrán ser descargados aquí:

- https://drive.google.com/file/d/12SXnHU-boiW_GDGGNLhr1BP9CsOsDhCL/view?usp=sharing

Incluye un archivo de tipo CSV:

- **Admission_Predict_Ver1.1.csv:**

Se trata de un archivo de CSV, en donde cada columna representa información del perfil académico de un estudiante dado, como por ejemplo el puntaje obtenido en un examen TOEFL, o la fuerza de una carta de recomendación. La consigna es tratar de predecir las posibilidades que tiene de entrar en un programa de Maestrías. (Se trata de estudiantes de la India)

Diccionario de datos:

1. GRE (General test) Scores (out of 340)
2. TOEFL (Examen de inglés) Scores (out of 120)
3. University Rating (out of 5)
4. SOP: Statement of Purpose (out of 5)
5. LOR: Letter of Recommendation Strength (out of 5)
6. Undergraduate GPA (Grade Point Average) (out of 10)
7. Research Experience (either 0 or 1)

8. Chance of Admit (ranging from 0 to 1)

Transformación del conjunto de datos

En este caso, la variable 8 es la variable dependiente. Y la que deberá ser calculada. Siendo que es una variable numérica. El problema es de regresión.

Se propone cambiar la variable 8, por una variable de tipo “etiqueta semantica”, de forma de cambiar el problema a un problema de clasificación. La clasificación sería la siguiente:

Rango	Etiqueta
$0.9 < X < 1$	Very_good
$0.8 < x < 0.9$	Good
$0.65 < x < 0.8$	Medium
$0.5 < x < 0.65$	Bad
$0 < x < 0.5$	Very_bad