

PROYECTO FINAL

INTRODUCCION A DEEP LEARNING

Fecha de entrega (presentación y código comentado): Lunes 6 de junio.

Fecha de presentaciones: Martes 7, miércoles 8 y jueves 9 de junio, de 18 a 19:30.

El proyecto consistirá en desarrollar un modelo de *Deep Learning*. Para la presentación de este trabajo se tomarán en cuenta los siguientes puntos:

- Muestreo de los datos.
- Proceso de definición de la topología de la red (pruebas que hayan hecho).
- Funciones de pérdidas y métricas utilizadas.
- Optimizadores.
- Regularizadores y drop-outs.
- Learning rates.
- Callbacks de Keras.
- Problemas con el sobreajuste (*overfitting*) y cómo los han solventado.
- Justificar el uso o no de *transfer learning*.
- Justificar la necesidad o no de *data augmentation*.

Los proyectos a elegir, con sus respectivas bases de datos, son los siguientes:

1. Determinación del género de una persona entre 25 y 35 años

https://drive.google.com/drive/folders/1FMlacOKczDwvIDL7JuY_jCxDdU7rU2iD

2. Identificación de la raza de perros entre:

Bichón maltés (n02085936-Maltese_dog)

Gran pirineo (n02111500-Great_Pyrenees)

Basenji (n02110806-basenji)

Lebrel afgano (n02088094-Afghan_hound)

<https://www.kaggle.com/jessicali9530/stanford-dogs-dataset>

3. Identificación de piedras entre:

Biotita, malaquita, cuarzo y bornita

<https://www.kaggle.com/asiedubrempong/minerals-identification-dataset?select=mineral>

4. Identificación de qué tan similares son dos textos que pueden ser diferentes por abreviaciones, textos incompletos, etc.:

<https://www.kaggle.com/rishisankineni/text-similarity>

5. Diferenciar entre tweets neutrales, positivos y extremadamente negativos con respecto al COVID-19:

<https://www.kaggle.com/datatattle/covid-19-nlp-text-classification>

6. Identificación de textos de acuerdo a su área científica, física, biología o química:

<https://www.kaggle.com/vivmankar/physics-vs-chemistry-vs-biology>

7. Detección de insuficiencia cardiaca

Predecir si una persona puede fallecer por una insuficiencia cardiaca, considerando diversos factores de salud, sus hábitos y otros factores descritos en la siguiente base de datos:

<https://archive.ics.uci.edu/ml/datasets/Heart+failure+clinical+records#>

8. Clasificación de género de películas entre:

Science fiction, romantic comedy, crime drama, musical.

<https://www.kaggle.com/datasets/jrobischon/wikipedia-movie-plots>

9. Clasificación de género de películas entre:

Crime, Western, Romance, Thriller

<https://www.kaggle.com/datasets/jrobischon/wikipedia-movie-plots>

10. Clasificación de género de películas entre:

Action, Horror, Comedy, Drama.

<https://www.kaggle.com/datasets/jrobischon/wikipedia-movie-plots>

Referencias:

F. Chollet, Deep Learning with Python, 1a y 2a edición.