Assignment 4: Deep Learning

Submission: Friday July 5th
3 students per group

Prof. Fabio A. González
Machine Learning - 2019-I
Maestría en Ing. de Sistemas y Computación

1. Design a function able to find the missing characters from a word. The function must work as follows:

```
>>> hangman("pe_p_e")
'people'
>>> hangman("phi__sop_y")
'philosophy'
>>> hangman("si_nif_c_nc_")
'significance'
>>> hangman("kn__l_d_e")
'knowledge'
>>> hangman("inte_r_ga_i_n")
'interrogation'
```

The function must be able to deal with up to 4 unknowns in arbitrary length words. The function must work in a reasonable time (max 30 seconds in a laptop).

2. **Movie poster classification**

   (a) Direct prediction

      i. Download the dataset MM-IMDB from `http://lisi1.unal.edu.co/mmimdb/`. The dataset includes poster images from movies of different genres. Find the six most frequent genres, chose images only from those genres, filter the images that only belong to one genre. Report the size of the resulting training, development and test datasets per genre.

      ii. Use Keras and the MobileNet pre-trained model, to classify the images in the dataset. Construct a confusion matrix that relates the six genres classes with the 10 most frequent classes from ImageNet predicted by the model.

   (b) Transfer learning

      i. Use the pre-trained MobileNet model as a feature extractor. Create a new model that replaces the top part of MobileNet with two layers of 256 and 6 neurons respectively.

      ii. Train the model with the training images from the bird dataset.

      iii. Evaluate the performance over the test dataset reporting the results in a confusion matrix. Discuss the results.

   (c) Fine tuning

      i. Repeat the experiment from the last question, but this time allow all the layers to be trained.

ii. Compare and discuss the results.

3. The assignment must be submitted as a <u>Jupyter notebook</u> through the following <u>Dropbox file request</u>, before midnight of the deadline date. The file must be named as `ml-assign4-unalusername1-unalusername2-unalusername3.ipynb`, where `unalusername` is the user name assigned by the university (include the usernames of all the members of the group). In case you need to include supporting files in addition to the notebook, submit a zipped file containing all the files and the notebook. Make sure that the notebook renders correctly and is free of errors before submitting.