Intro to DL Y2022 Semester B

Project part 3

Submission

Submission is in pairs or singles.

Data format and description

The data appear in moodle in three files:

- training_ex3_dl2022b.csv
- validation_ex3_dl2022b.csv
- test_ex3_dl2022b.csv

Your data are images linked to in given csv files, and your labels are the names of what's in them. The test file comes without captions.

The task

Your goal is to **predict the label** for test data values, based on the image (do not use the ID! If you do, your score will be 0, and you will be on the bottom of the kaggle table).

1. Define a CNN network that will receive images and output the word vector of their caption (average of word vectors for all words in the caption).
2. Train your model on the training data, assess on the validation set. Then run your best model on the test set to get predictions.

Training/validation data format

id	image	label
1	http://...	apple
2	...	...
...	...	...

3. Define a CNN2D or a CNN2D model keras.
   a. Use any number of layers you want.
   b. The last layer should have # of neurons=# of labels, your loss should be **categorical crossentropy** (https://www.tensorflow.org/api_docs/python/tf/keras/losses/CosineSimilarity), and your metric should be **accuracy** as well (https://www.tensorflow.org/api_docs/python/tf/keras/metrics/CosineSimilarity).
4. Evaluate your model on the test set and save the label you produce for every image in .csv file, as follows:
   a. ID=image ID

b. label=text label (the name of the fruit).

Note: you can use any of the normalization and data analysis methods in sklearn to improve your scores.

Result submission

Your result should include item IDs from the test set and predicted label, and to be saved as csv file:

id	label
1	apple
2	apple
…	apple

How kaggle works (a reminder)

Your results will be compared with the actual test dataset labels, and the resulting accuracy will be reported on the scoreboard of the competition. Note that public scoreboard will show accuracy on 50% of the test set, and private (i.e., my) scoreboard will show accuracy on the whole test set. The final scoreboard will be published after submission & code checking is over, and your grade will be determined by your place in the competition.

Code submission

Submit your code on moodle, as a single <id1>_<id2>.py file (do not submit python notebooks!).

Note of warning: all code will be automatically checked for copying. If cheating is discovered, you will get grade 0 automatically and go on to face the scholarly committee.

How to boost your results (ideas):

1. Larger images = better accuracy, but slower network.
2. Can you use 2 networks? Maybe some fruits are easier to identify?