ARTIFICIAL INTELLIGENCE (E016330)
ALEKSANDRA PIZURICA
GHENT UNIVERSITY
2019/2020 AY

Assistants: Laurens Meeus, Nina Žižakić
Professor: Aleksandra Pižurica
E-mail: ai@lists.ugent.be

Practical Session 3: Neural networks

1 Introduction

In this practical session, in collaboration with Robovision, we will use *Keras* [1] and *Google Colab* [2] to get familiar with different theoretical and practical aspects of Neural Networks.

Google Colab is a free cloud service to run Python notebooks with GPU support, ideal for taking our first steps in deep learning.

In preparation for this practical session, take a look at *Google Colab* and *Keras*:

- https://colab.research.google.com/notebooks/welcome.ipynb
- https://keras.io/

As training a network can easily take minutes to complete, we advise you to take a look at other questions while training a model in the background.

2 Database

The assignment will be on the classification of fruits and vegetables. The data is obtained from Horea Muresan's "Fruit recognition from images using deep learning" [3]. The "Fruits-360" contains 75,937 images of 120 fruits. For this practical session, we will work on a subset, already dividing the training, validation and test set. It will not be necessary to download the dataset locally.

- 175MB dataset

3 Practical tasks and questions

We will start from a provided notebook *ai_practicals_nn.ipynb*.

- https://drive.google.com/open?id=1m2qp-LRaOMbxcsvQlowXusy4WsYZ_NRQ

Copy the notebook to your own Google drive to freely make modifications. The notebook will act as a guidance to answer the different questions. You are free to make any adjustments to the script.

3.1 Loading data and preprocessing

Run the notebook up to and including section **Database**.

1. Take a look at the input data x.
   (a) What is the amount of training, validation and test samples?
   (b) What is the size of each image?
   (c) Plot some of the images. If we would add new images to the dataset, what preprocessing would be necessary?
2. Take a look at the output data y.
   (a) What are the different classes?
   (b) What is the meaning of each sample of y?
   (c) For the training, validation and test set: how much samples are there of each fruit type?

3.2 Training

3.2.1 Apple classification

We'll start with a simple binary classification: Is a fruit an apple or not.

3. Train the provided network and study the generated logs using Tensorboard. You might have to kill the previous Tensorboard in order for it to update.
   (a) Keep training the model untill the training loss starts to converge.
   (b) Show the training and validation loss for each epoch (you may ignore the test set for now).
   (c) By looking at both loss curves, do we speak of generalization or specialization? Explain.

Besides accuracy, other metrics like precision and recall, are often used to measure the classification performance.

4. Generate the prediction of the test set.
   (a) What does this prediction output represent?
   (b) How do you convert the output of the model to a binary prediction of apple or not-apple?
   (c) In a table, show the confusion matrix of the test set using the ground truth and prediction.
   (d) The F1-score is another score often used instead of accuracy. Look up the definition of the F1-score and explain why in many cases the F1-score is a better measure of performance than the accuracy. What score would you advise to use to assess our Apple classifier?

(e) Receiver Operating Characteristic (ROC) plots the precision in function of recall depending on the decisions threshold. Show the ROC curve of our model and explain what you see.

The previous assignments were done on a prefixed model, however to find the optimal network architecture for an application, numerous hyperparameters have to be correctly chosen:

- Number of layers
- Number of filters per layer
- Loss function
- Learning rate
- ⋯

To optimize the learning rate, we will train a model with a varying learning rate from very low to very high.

5. Run the cell that trains the network with an incremental increasing learning rate.
   (a) Explain what the learning rate is.
   (b) Based on the results, which learning rate would you propose to use and why?
   (c) What happens when you set the learning rate too low?
   (d) What happens when you set the learning rate too high?

Secondly, we will test the performance of different models with a varying amount of feature-maps. For this, take a look at the method `get_model_apple_simple`.

6. Based on the provided model, select a single layer of the model and train multiple models with each a varying number of features for this layer. You can leave the rest of the model the same.
   (a) Which of your models has the best performance? Explain how you assessed the performance of the models.
   (b) For your best model, run `<model_name>.summary()`. Explain how the number of parameters per layer are calculated.
   (c) In general, what are the disadvantages of a layer that is too narrow (low number of feature-maps) or too wide (high number of feature-maps). Use terms like underfitting, specialisation and generalization.

3.2.2 Fruit classification

Before now the network was trained to distinguish between apples and not-apples. We will now alter the model to distinguish between the different kinds of fruits and vegetables.

7. Take a look at the code where the previous model is build. Make now your own model that can classify multiple classes:

(a) The amount of filters of the last layer of the model should be changed from 1 to the amount of classes there are.

(b) Replace the activation function of this last layer from sigmoid to softmax. Explain the difference and thus explain why in our case we can use the softmax function.

(c) Replace the loss-function from mean-squared-error (mse) to categorical_crossentropy (not te be confused with binary crossentropy). Just like accuracy vs F1-score, explain why crossentropy is preferred over mse.

8. Train your network untill it convergences or untill the model starts to overtrain.

   (a) Show the loss and accuracy with respect to number of epochs.

   (b) Using the test set, predict the classes of the test set samples and show the confusion matrix with all the classes. Analyze which fruits are difficult to distinguish by the model.

3.3 Testing

By now you should have models to classify between apples and not-apples and to distinguish between different kind of fruits.

9. Test the models on your own set of pictures. Consider that some form of preprocessing might be needed before the model can correctly handle it (see question 1).

   (a) Provide examples of a True Positive, True Negative, False Positive and False Negative.

   (b) Are our trained networks ready for real-life applications? If not, what modification would you suggest are needed to make the biggest improvements.

4 Submission

The project is done in groups of maximum two students. Submit your report of the practice together with your notebook.

Important remarks:

- Write concise answers to the questions. Indicate clearly the number of each task/question when answering it. Subquestions may be combined if it improves readability.
- The overall score may be influenced by the report presentation too (pay attention to the clarity of answers and avoid sloppiness).
- The submitted file should be named after both authors and specify the project subject (e.g. Frodo Baggins & Samwise Gamgee: nn-fbaggins-sgamgee.zip).
- It is the student's responsibility to check that the submitted documents are complete and not corrupted. Avoid any situation that prevents the correct evaluation of your report.
- Deadline: **December 13, 2019, 23:59 CET**

References

[1] F. Chollet *et al.*, "Keras." https://keras.io, 2019.

[2] Google, "Welcome to colaboratory!." https://colab.research.google.com/notebooks/welcome.ipynb, 2019.

[3] H. Muresan, "Fruit recognition from images using deep learning," *Acta Universitatis Sapientiae, Informatica*, vol. 10, pp. 26–42, 06 2018.