

MSc - Media and Textmining

Homework results

Daniel Mark Kiss

2023

# Contents

|                  |   |               |
|------------------|---|---------------|
| <b>Chapter 1</b> | <b>Task description</b>   | <b>Page 2</b> |
| <b>Chapter 2</b> | <b>Elaborated work</b>  | <b>Page 3</b> |
| 2.1              | program.py  | 3             |
| 2.2              | cnn.py  | 4             |
|                  | init — 4 • split — 5 • compile — 5 • train — 5 • evaluate — 5 • predict — 5 |               |
| 2.3              | labeler.py  | 5             |
| 2.4              | filereader.py   | 5             |
|                  | read_training_files — 6 • create_train_set — 6                              |               |
| 2.5              | Tensorflow using GPU  | 6             |
| <b>Chapter 3</b> | <b>Results and parameter optimization</b>                                   | <b>Page 7</b> |

# Chapter 1

## Task description

The data set contains small black and white images labelled with character identifier. The task is to recognize the character (0-9, a-z, A-Z) of unknown images in the test set by deep learning classifier without any human activity (human's help is equivalent to cheating). Work out a deep learning classifier model, and train this model by the training dataset! Use cross-validation in order to measure the goodness indicators (accuracy, AUC) of your model! Please pay attention to the parameter optimization! Please investigate the model, the result and describe the details of your solution!

# Chapter 2

## Elaborated work

In the chapter I will present the elaborated work that I have done. For my work I have created separate classes and files for the different subtasks of the Homework. In every section I will present and explain what the different classes are used for.

### 2.1 program.py

Program.py is the core of the program. It is the main file that is used to run the program. It is responsible for the following tasks:

1. Initializing the FileReader class
2. Running the FileReader class functions
  - read\_training\_files()
  - create\_train\_set()
3. Initializing the ConvolutionalNeuralNetwork class
4. Running the ConvolutionalNeuralNetwork class functions
  - split()
  - compile()
  - train()
  - evaluate()
5. Reading the test files
6. Creating the test set

## 7. Exporting results to output.txt file

When we would like to run the program we should also import the `FileReader` and `ConvolutionalNeuralNetwork` classes. For running we should type `python program.py` to the terminal and the program should start running.

## 2.2 `cnn.py`

This file contains the fundamental code for the neural network. I used many other libraries for my homework like *numpy*, *tensorflow*, and *sklearn* also.

### 2.2.1 `init`

In the `__init__` function I initialize the arrays which I used for training and splitting. These are the different variables that I used in my class:

- `self.images`: contains the raw images
- `self.label`: contains the labels for the images
- `self.img_train`: contains the train set of images
- `self.img_test`: contains the test set of images.
- `self.label_train`: contains the labels of the train set
- `self.label_test`: contains the labels of the test set
- `self.model`: is a model used from `tensorflow.keras`. It has five layers, three 2D convolutional layers and two 2D MaxPooling layers. The first convolutional layer has 32 filters/kernels. Each filter is 3x3 matrix. I choose ReLU as the activation function. The input shape is 128x128, 1 which is set because the given images are this size and they are grey scaled. The MaxPooling layers are used for down-sampling and reducing the spatial dimensions of the input. It uses a 2x2 pooling window/filter. It also helps reducing the chance of overfitting. It only retains only the most important information from it.

After the Initialization of the model I add a Flatten layer to it. It is used for converting and outputting a 1D array. It is necessary before passing to a dense, fully connected layer. The next line adds a layer with 64 units (neurons) and a ReLU activation function. A dense layer means that each neuron in this layer is connected to every neuron in the previous layer. The last line in the `init` function adds another dense layer with 62 units, representing the output layer. The activation function used in the output layer is "softmax." Softmax is often used for multi-class classification problems. It

converts the raw output scores into probabilities, where each value represents the probability of the input belonging to a particular class. The sum of all probabilities for a given input is 1, making it suitable for classification tasks like this task also.

### **2.2.2 split**

The split function is used for splitting the images and the labels into training and testing set, with the test set is 33% of the input. I also add a random\_state seed so it will result always the same output.

### **2.2.3 compile**

The compile function is used for compiling the model. I used the Adam optimizer for the compilation. It is an adaptive learning rate optimization algorithm that's been designed specifically for training deep neural networks.

### **2.2.4 train**

The train function is used for training the model. I used 15 epochs for the training. I also added a LearningRateScheduler to the training. It is used for changing the learning rate during the training. I used 64 as the batch size. The batch size is a number of samples processed before the model is updated. The validation split is set to 0.2. It means that 20% of the training data is used for validation.

### **2.2.5 evaluate**

The evaluate function is used for evaluating the model. It returns the loss and the accuracy of the model.

### **2.2.6 predict**

The predict function is used for predicting the labels of the test set. It returns the predicted labels. It only processes one image and gives a prediction for it. It is used for the output.txt file.

## **2.3 labeler.py**

The labeler.py file is used for labeling the images. It is only a huge dictionary with the labels for the images.

## **2.4 filereader.py**

The filereader.py file is used for reading the images and the labels from the given files.

### 2.4.1 read\_training\_files

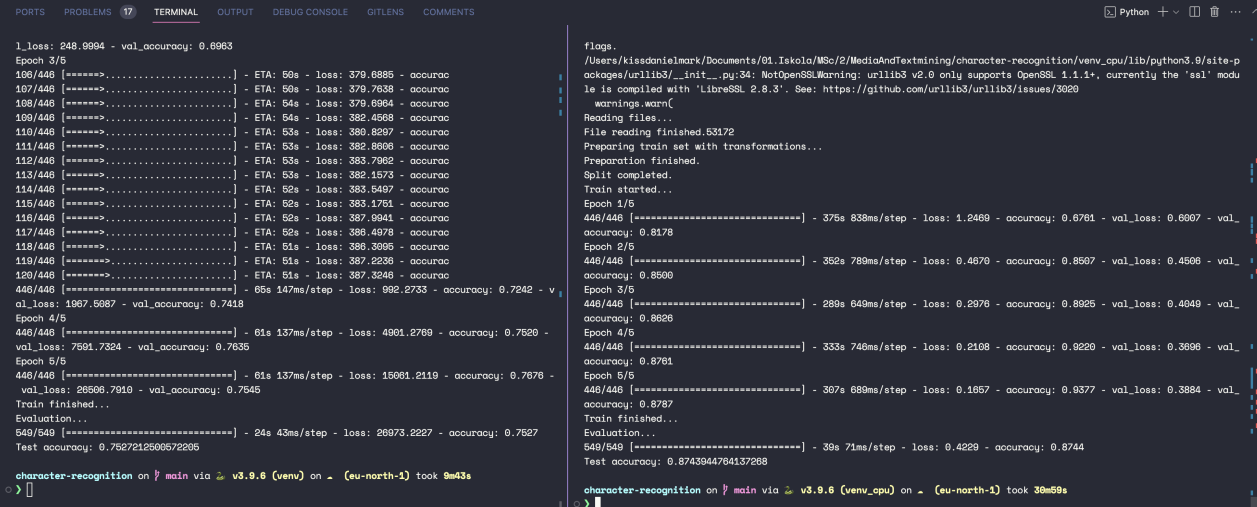
It reads the folders and images recursively. These functions also used for reading not only the training folders but the test folder as well.

### 2.4.2 create\_train\_set

For every image that has been read it is processed by a private function called `_process_image()`. These processed images are added to the images array and the labels are added to the labels array. The `_process_image()` function is used for creating the labels for the images extracted from the filename. Also the images are converted to numpy arrays. After that the function is normalizing the images.

## 2.5 Tensorflow using GPU

I used tensorflow-gpu for my homework. It is a version of tensorflow that uses the GPU for the calculations. I had an interesting problem with it. When I ran the same code on CPU and GPU I got different results. On CPU the model gave a 87% accuracy, but on GPU it gave a 75% accuracy.



```
l_loss: 248.8994 - val_accuracy: 0.6963
Epoch 3/5
106/446 [====>.....] - ETA: 59s - loss: 379.6885 - accurac
107/446 [====>.....] - ETA: 59s - loss: 379.7638 - accurac
108/446 [====>.....] - ETA: 54s - loss: 379.6964 - accurac
109/446 [====>.....] - ETA: 54s - loss: 382.4568 - accurac
110/446 [====>.....] - ETA: 53s - loss: 380.8297 - accurac
111/446 [====>.....] - ETA: 53s - loss: 382.8008 - accurac
112/446 [====>.....] - ETA: 53s - loss: 383.7862 - accurac
113/446 [====>.....] - ETA: 53s - loss: 382.1573 - accurac
114/446 [====>.....] - ETA: 52s - loss: 383.5497 - accurac
115/446 [====>.....] - ETA: 52s - loss: 383.1761 - accurac
116/446 [====>.....] - ETA: 52s - loss: 387.0941 - accurac
117/446 [====>.....] - ETA: 52s - loss: 386.4978 - accurac
118/446 [====>.....] - ETA: 51s - loss: 386.3995 - accurac
119/446 [====>.....] - ETA: 51s - loss: 387.2238 - accurac
120/446 [====>.....] - ETA: 51s - loss: 387.3248 - accurac
446/446 [=====] - 65s 147ms/step - loss: 992.2733 - accuracy: 0.7242 - v
al_loss: 1967.5687 - val_accuracy: 0.7418
Epoch 4/5
446/446 [=====] - 61s 137ms/step - loss: 4901.2769 - accuracy: 0.7520 -
val_loss: 7691.7324 - val_accuracy: 0.7635
Epoch 5/5
446/446 [=====] - 61s 137ms/step - loss: 15061.2119 - accuracy: 0.7676 -
val_loss: 26506.7919 - val_accuracy: 0.7545
Train finished...
Evaluation...
549/549 [=====] - 24s 43ms/step - loss: 26973.2227 - accuracy: 0.7527
Test accuracy: 0.7527212590572205

character-recognition on / main via v3.9.6 (venv) on a (eu-north-1) took 9m43s

flags
/Users/kissdanielmark/Documents/G1.Iekola/WSs/2/MediaAndTextmining/character-recognition/venv_cpu/lib/python3.9/site-p
ackages/urllib3/_init_.py:34: NotOpenSSLWarning: urllib3 v2.0 only supports OpenSSL 1.1.1+, currently the 'ssl' modu
le is compiled with 'LibreSSL 2.8.3'. See: https://github.com/urllib3/urllib3/issues/3020
  warnings.warn(
Reading files...
File reading finished.53172
Preparing train set with transformations...
Preparation finished.
Split completed.
Train started...
Epoch 1/5
446/446 [=====] - 375s 838ms/step - loss: 1.2489 - accuracy: 0.6761 - val_loss: 0.6007 - val_
accuracy: 0.8178
Epoch 2/5
446/446 [=====] - 352s 780ms/step - loss: 0.4670 - accuracy: 0.8567 - val_loss: 0.4506 - val_
accuracy: 0.8500
Epoch 3/5
446/446 [=====] - 289s 649ms/step - loss: 0.2976 - accuracy: 0.8925 - val_loss: 0.4849 - val_
accuracy: 0.8626
Epoch 4/5
446/446 [=====] - 333s 746ms/step - loss: 0.2108 - accuracy: 0.9220 - val_loss: 0.3698 - val_
accuracy: 0.8761
Epoch 5/5
446/446 [=====] - 307s 689ms/step - loss: 0.1657 - accuracy: 0.9377 - val_loss: 0.3884 - val_
accuracy: 0.8787
Train finished...
Evaluation...
549/549 [=====] - 39s 71ms/step - loss: 0.4229 - accuracy: 0.8744
Test accuracy: 0.8743944764137268

character-recognition on / main via v3.9.6 (venv_cpu) on a (eu-north-1) took 39m59s
```

I tried to solve this problem but I could not find the solution for it.

## Chapter 3

# Results and parameter optimization

In this chapter I will showcase the results of my homework and the parameter optimization that I have done.