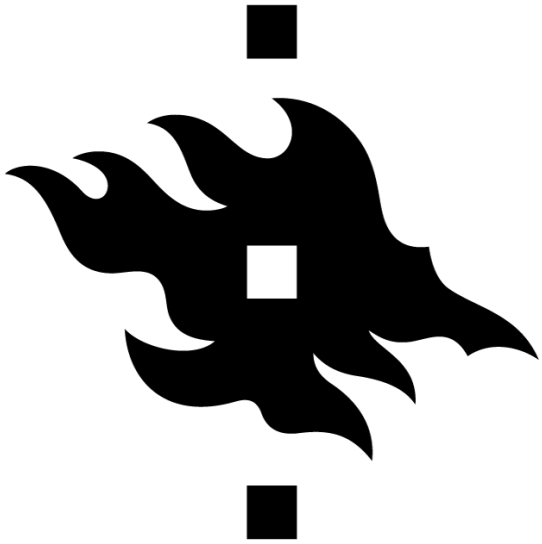
University of Helsinki



Report of the project of Deep Learning

Image project

Students:
Riccardo Menoli
Stefano Lia
Ahmed

2018/2019

Contents

1 About the project

2 Data Management

To start the project the first thing to do was finding the proper way to handle the data. In this section we explain all the problems we faced and the corresponding found solution during this process.

2.1 Creation of the Dataset

The first problem was the **creation of the Dataset**. As the guide of PyTorch and the exercise suggested we created a custom Dataset following this guide. Based on this guide we organized the pandas Dataframe in the following way:

- one column containing the path of the image;
- one column for each class. If the sample (the row) belongs to one on more class, it has 1 in the corresponding column which represents the belonging class;
- one column containing the representing image in the PIL format;
- one column containg one list whose values are the classes which the example belongs to.

2.2 Grayscale problem

2.3 Missing values problem

2.4 Imbalanced data

3 Models

4 Conclusion