TBMI26 – Computer Assignment Reports
Deep Learning

Deadline – March 14 2021

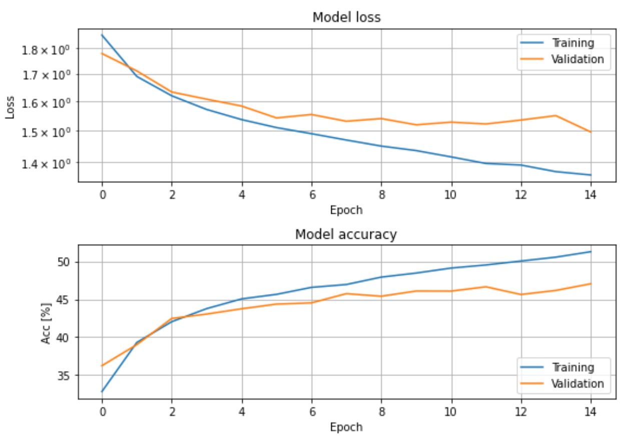
Author/-s: Niclas Hansson: nicha207
Måns Aronsson: manar189

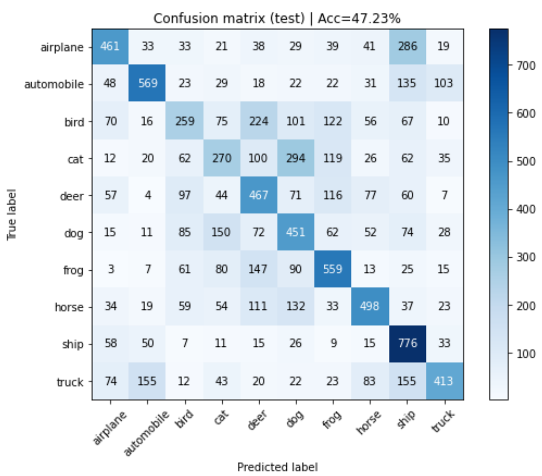
In order to pass the assignment you will need to answer the following questions and upload the document to LISAM. Please upload the document in PDF format. **You will also need to upload the Jupyter notebook as an HTML-file (using the notebook menu: File -> Export Notebook As...)**. We will correct the reports continuously so feel free to send them as soon as possible. If you meet the deadline you will have the lab part of the course reported in LADOK together with the exam. If not, you'll get the lab part reported during the re-exam period.

Disclaimer from the authors: We forgot to save the html file during the lab. The file provided with this report is the same file used to achieve the images for this report, however the results may differ slightly since we needed to rerun the code and save it again.

1. **The shape of X_train and X_test has 4 values. What do each of these represent?**
   In order: Number of images, pixel size of the images, image dimension (3 is color.)

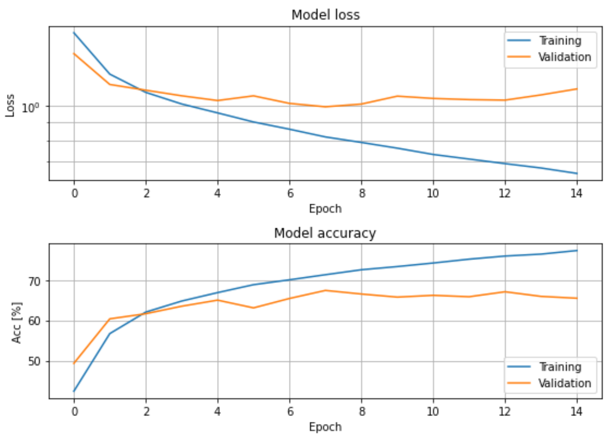
2. **Train a Fully Connected model that achieves above 45% accuracy on the test data. Provide a short description of your model and show the evaluation image.**
   The model uses two dense layers with a rectified linear unit as an activation-function. It uses >600 000 trainable parameters. The result does not improve with more training after reaching a cap on the validation data around 46-47%.

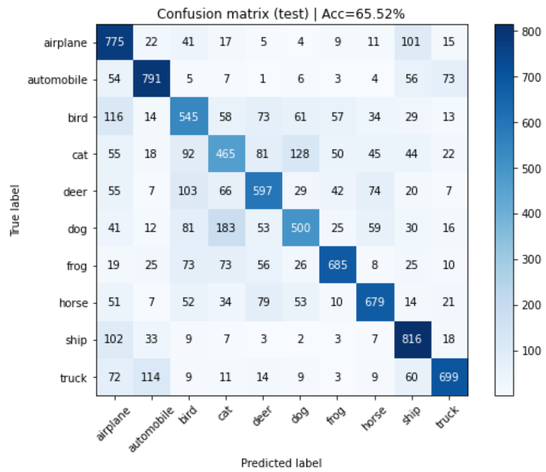




3. **Compare the model from Q2 to the one you used for the MNIST dataset in the first assignment, in terms of size and test accuracy. Why do you think this dataset is much harder to classify than the MNIST handwritten digits?**
   Numbers always have the same orientation. In this dataset a cat in one image can be angled one way and another cat can be its mirror image, meaning that the network must train more variables to identify objects and cannot heavily rely on a few defining features.

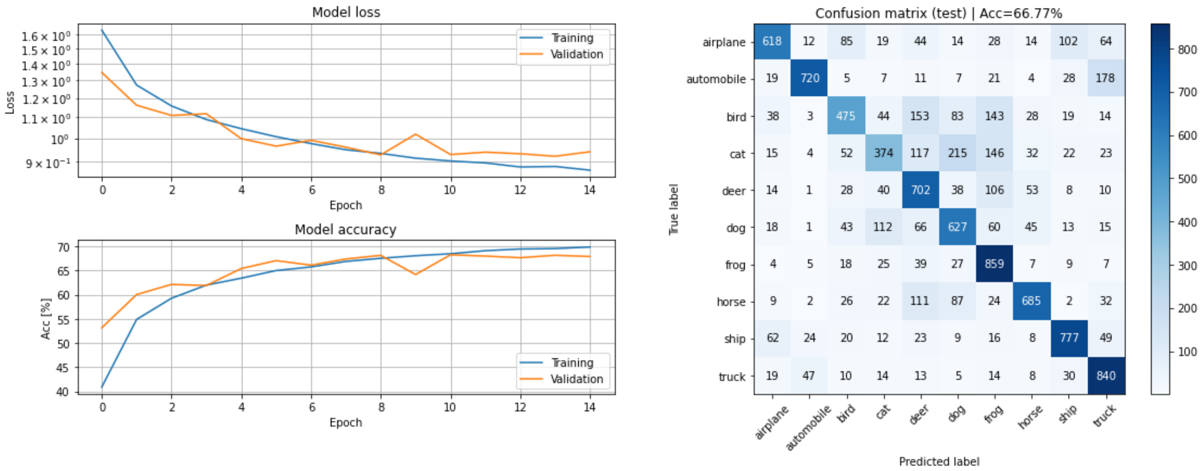
4. **Train a CNN model that achieves at least 62% test accuracy. Provide a short description of your model and show the evaluation image.**
   The CNN uses 2 blocks, where each block consists of a convolution with a 3x3 kernel and a pooling. The pooling down samples the image to half the size. The network uses about 40’000 trainable parameters but performs almost 50% better than just the dense layers.





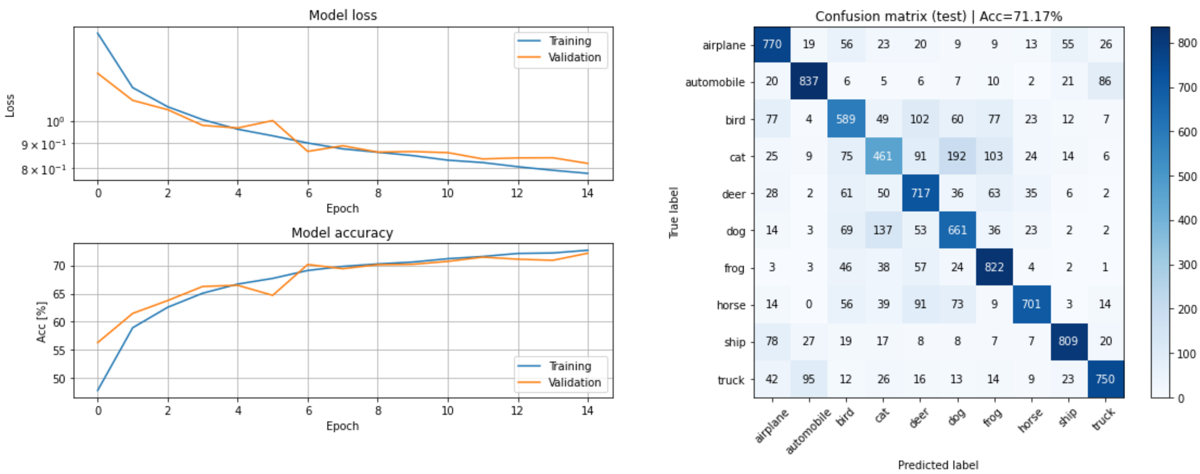
5. **Compare the CNN model with the previous Fully Connected model. You should find that the CNN is much more efficient, i.e. achieves higher accuracy with fewer parameters. Explain in your own words how this is possible.**
   The dense network uses about 14 times more variables but scores 50% lower. The convolutional network learns more information about each image due to the convolution and down sampling, meaning that it takes fewer variables to make a more well-informed decision.

6. **Train the CNN-model with added Dropout layers. Describe your changes and show the evaluation image.**
   Added a dropout-layer after each pooling which discards the output of 20% of the variables.



7. **Compare the models from Q4 and Q6 in terms of the training accuracy, validation accuracy, and test accuracy. Explain the similarities and differences (remember that the only difference between the models should be the addition of Dropout layers).**
   Hint: what does the dropout layer do at test time?
   The models perform similarly in terms of time and accuracy. However, the dropout-models generalizes much better, there is no overtraining since the dropouts forces the network to not rely on a few strong variables.

8. **Train the CNN model with added BatchNorm layers and show the evaluation image.**



9. **When using BatchNorm one must take care to select a good minibatch size. Describe what problems might arise if the wrong minibatch size is used.**
   You can reason about this given the description of BatchNorm in the Notebook, or you can search for the information in other sources. Do not forget to provide links to the sources if you do!

   Since BatchNorm calculates mean and variance for each batch the reliability of the result decreases if the batches become too small, this may cause the network to become unstable. Too big batch sizes do not generalize well.

https://towardsdatascience.com/curse-of-batch-normalization-8e6dd20bc304#_=_

10. Design and train a model that achieves at least 75% test accuracy in at most 25 epochs. Explain your model and motivate the design choices you have made and show the evaluation image.

We tried to implement the ResNet model from lecture 4. It uses blocks where most blocks consist of convolutions, normalizations and two relu activation function, one in the middle. The input to a block is added to the result of the block before the last activation function.

The network is trying to estimate a function and the feedforward within each block can be seen as F(x). This means that each consecutive block only needs to estimate the residual F(x)-x since x is added after the residual layer. The network will act as a shallower network during the early stages of the training process and as a deeper network in the later stages.