Technische Hochschule Ingolstadt, M-APE, Prof. A. Schiendorfer Lukas Lodes M.Sc.

Intro to AI and Neural Networks (Summer 2023)

Assignment 10

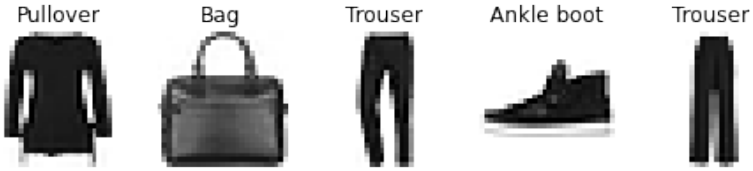
Exercise 1 ***(Activation functions)***

In `Activation_Functions.ipynb` you'll find a notebook exploring the characteristics of different activation functions.

- Take a look at the implementations of the differente activation functions. How does ELU (which we have not covered in the lecture) differ?
- What gradient behavior do you observe? Which activation functions look acceptable from a gradient perspective?

Exercise 2 ***(Optimizing hyperparameters for the classifier on fashion MNIST)***

We will revisit last exercise's classifier and explore how different activation functions and learning rates affect the training performance. Next, we will again train neural network on the fashion MNIST dataset. This dataset contains 28×28-images of items of clothing (and is slightly more challenging than the regular MNIST dataset for handwritten digit recognition):



This exercise will require Keras[1]. Data loading and preparation of train/test/val splits have been prepared in `Fashion_MNIST_hyperparams_for_students.ipynb`.

- Try the ReLU activation function, a leaky ReLU variant, sigmoid and tanh for about 30 episodes each. Plot the different learning curves. Which one converges fast, which one reaches a good accuracy?
- Experiment with different learning rates for your chosen activation function. Explore values in $\{0.1, 0.01, 0.001\}$.
- Once you've settled on a suitable activation function and learning rate. Train the model with these hyperparameters for 50 episodes. Evaluate train, test, and val accuracy.

[1] `keras.io` – if you want to run it locally, you need to install TensorFlow `https://www.tensorflow.org/install` or just use Google Colab.