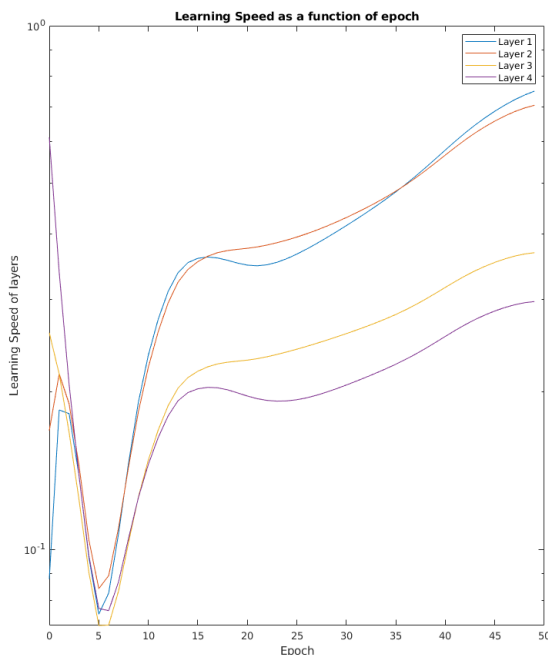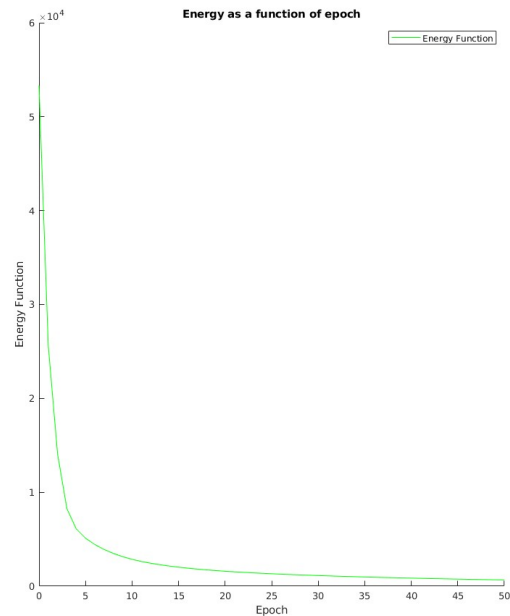# Artificial Neural Networks

Homework 3 Exercise 2

Plot of the Energy Function per Epoch

It's clear that the energy function drops drastically first around epoch 5, and then continues to drop to a lesser degree in the rest of the epochs, the change appears to be logarithmic.

Plot of the Learning Speed per Epoch

In the initial phase of the training, the learning speed drops sharply, when the initial weights are just random numbers, showing the vanishing gradient problem and how it affects the learning process. However after the weights are no longer random the network starts to gradually learn and adjust to the training patterns, yielding relatively higher learning speed for different layers. Deep into the learning process, layers don't change their weights randomly, because they start to specialize in feature-detecting/boundary-drawing.