TU München
Fakultät für Informatik
PD Dr. Rudolph Triebel
J. Chiotellis, C. Hazirbas, P. Haeusser

Machine Learning for Robotics and Computer Vision
Winter term 2015

Homework Assignment 3
Topic: Deep Learning
November 23th, 2015

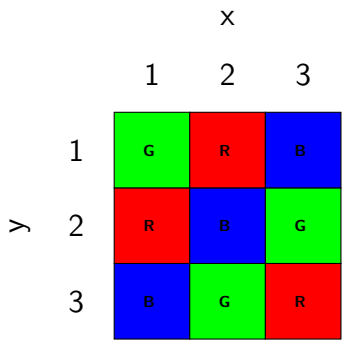
Hidden Markov Models and Mixture Models

Exercise 1: Viterbi Algorithm

We play again with our robot from the first homework assignment. As we mentioned back then the robot has a camera with an observation model that looks as follows:

Sensed color \ Actual color	R	G	B
R	0.8	0.1	0.1
G	0.1	0.6	0.2
B	0.1	0.3	0.7

This time we put the robot in a room where the floor looks like this:



a) What is the state space? What is the observation space? Draw the trellis diagram.

b) Assume the robot can only move vertically and horizontally. We let the robot move randomly. If the attempted move leads outside of the bounds of the room the robot stays at its current position. Compute the state transition matrix.

c) After 3 time steps, what is most likely the path that the robot followed if the camera reads $\{z_1 = \text{R}, z_2 = \text{G}, z_3 = \text{G}\}$? Assume the robot's initial position is unknown.

d) Solve the exercise in your preferred programming language.

Exercise 2: K-Means and Expectation-Maximization (Programming)

Download the fisher-iris.zip file. Inside you will find Fisher's Iris dataset. It consists of 150 flowers that have 4 properties and belong to 3 different species. Implement following exercises in your preferred programming language.

a) Implement the K-Means algorithm to cluster the dataset.

b) Run your algorithm 100 times with randomized initial values for the cluster centers. Each time compare the result with the true labels. How much does the performance change for different initializations?

c) Now implement the Expectation-Maximization algorithm (EM) for Gaussian mixture models (GMM) using a fixed symmetric covariance (supplied as a parameter) and apply it to the same dataset.

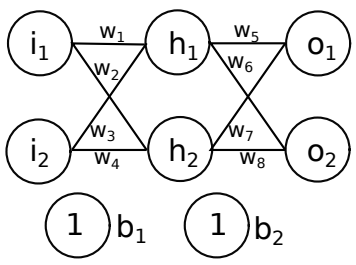
d) Again evaluate the algorithm for different initializations of cluster centers and covariances. How do the EM errors compare to K-means on average? Does EM converge faster? Is it more reliable in finding the true means?

Neural Networks and Deep Learning

Please read `http://neuralnetworksanddeeplearning.com/chap1.html` and `chap2.html`. You don't have to implement the network yourself nor do the exercises. Please first complete the reading above. We are going to follow Michael Nielsen's notation.

Exercise 3: Back Propagation

- Suppose we modify a single neuron in a feedforward network so that the output from the neuron is given by $f(\sum_j w_j x_j + b)$, where f is some function other than the sigmoid. How should we modify the backpropagation algorithm (from chapter 2 of the above reading) in this case?
- Compute the gradient of the cost function C respect to w_5 $\left(\frac{\partial C}{\partial w_5}\right)$ given the following network:



$C = \sum_{i=1}^{2}(t_i - a_i)^2$ where t_i is the target value for the respective output neuron o_i and a_i is the output of the neuron o_i. Input/output of h_1, h_2, o_1, o_2 is computed as $z_i^l = \sum_j w_j \cdot a_j + b_l$, $a_i^l = \sigma(z_i^l)$ where $\sigma(\cdot)$ is any activation function.

Exercise 4: Convolutional Layer Arithmetic

Consider a very simple convolutional neural network that just consists of one convolutional layer. It has the following parameters:

- number of kernels: $num = 64$
- size of kernels: $k = 3 \times 5$
- stride: $s = 2$
- padding: $p = 1$

Assume, the input to this layer is an a batch of RGB images. There are 10 images in one batch and the images have a dimension of 123×81.

a) What is the shape of the input blob to the convolutional layer? Hint: it's a tensor with four axes.

b) What is the shape of the output blob of the convolutional layer?

The next exercise class will take place on **December 4th, 2015**.

For downloads of slides and of homework assignments and for further information on the course see

`https://vision.in.tum.de/teaching/ws2015/mlcv15`