Assignment 34

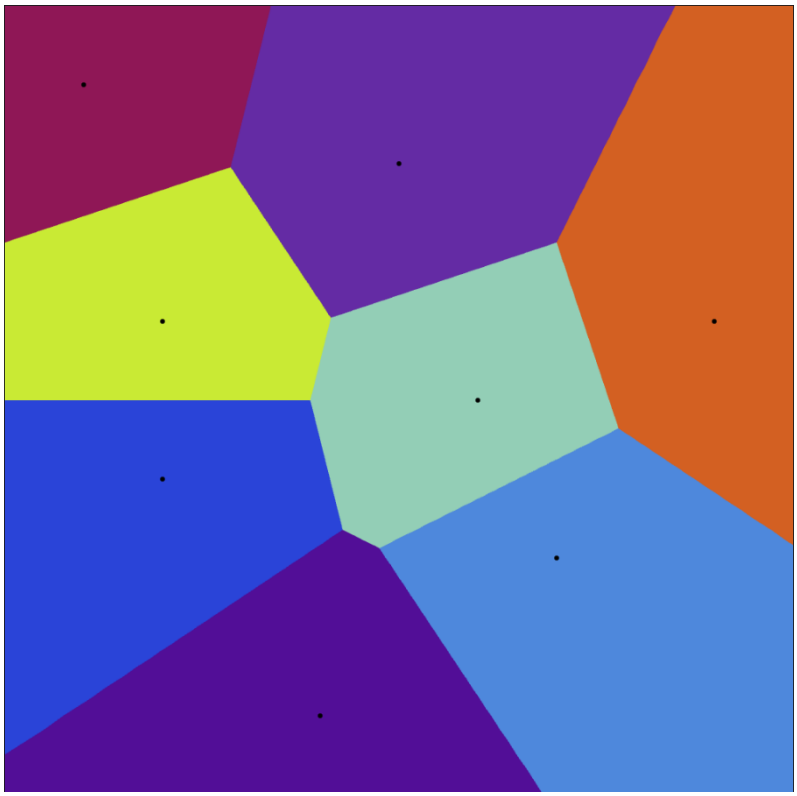
Exponential decay: n(t) = n_i*(n_f/n_i)^(t/t_max),
where n_i and n_f are initial and final values of the learning rate, and t_max is the total number of adaptation steps taken.

The figure below shows the decaying learning for a particular set of parameters: n_f = 1e-3, n_i=1, t_max = 10000, in logarithmic scale

Assignment 36

A neuron j could be the winner only if we give the same input pattern twice. If neuron i was the last winner, it would be if the learning rate n(t) or the neighborhood function h(dist(i,j), t) would be bigger than 1. That is not possible
$0 < n(t) < 1$
$0 <= h(dist(i,j), t) <= 1$

Assignment 38
Algorithm: The line segments of the Voronoi diagram are all the points in the plane that are equidistant to the two nearest sites. So, for each point, connect it with the closest points, draw a perpendicular line through the middle. The Voronoi vertices (nodes) are the points equidistant to three (or more) sites - the points of intersection of the perpendicular lines.



Assignment 33

The structure of a SOM consists of: A distance metric in input space. A distance metric dist() defined on the grid G. A neighborhood function h(dist), adjusting the amount of learning. A labeling procedure that can assign a meaning for every reference vector or center C from V. Steps of training a Self Organizing Feature Map (SOM):

1. Initialize SOM. Get some training data ${}^{p}X$ - P patterns from an area V of the N-dimensional input space, and labeling data. Define the grid structure, network size K. Initialize the center vectors C_k of each neuron.
2. Choose a stimulus and apply it to the SOM. Choose a pattern X which is called stimulus from V. All K neurons receive the same input vector X (stimulus).
3. Calculate the response. The applied pattern ${}^{p}X$ is compared to the respective center vector C_k of each SOM neuron k. The result of this comparison is called response r_k. If distance is Euclidean $r_k = \|C_k - {}^{P} X\|_2$.
4. Determine the winner i. The neuron i with the best response will be denoted the winner. $\|C_i - {}^{P} X\|_2 \leq \|C_j - {}^{P} X\|_2$ for $i, j = 1...k$.
5. Apply learning rule. Move the center C_j a bit towards the pattern ${}^{p}X$ if this neuron is close to winner C_i. $\Delta C_j = \eta(t)h(dist(i, j), t)({}^{p}X - C_j)$ where $0 < \eta < 1$ - learning rate decaying over time t, $h(dist, t)$ - neighborhood function (shrinking with time t), $dist(i, j)$ distance between neuron i and neuron j defined on the grid.
6. Update centers. The new values for the center vectors C_k will be calculated by adding the change ΔC_k in center vector to the center vector C_k. ${}_{new}C_k =_{old} C_k + \Delta C_k$.
7. Continue learning? If stopping condition is not satisfied go to step 2.
8. Labeling, assign a meaning to the neurons.

Assignment 34

Assignment 35

We will define a criterion n that measures the quality of the SOM as an average distance between data vector in input space X and its best match in V. It is also called expected quantization (or distortion) error.

$$E = \frac{1}{N} \sum ||^p X - C_i||$$

This way better mapping will yield better result because of smaller average distances.

Assignment 36

Assignment 37

Classification is a task of assigning information to predefined classes (supervised learning), whereas clustering is a task of dividing information based on features that are not given, but must be extracted from data (unsupervised learning). Other difference is that clustering works with data in hand, but classification is used to predict new data. In [1] we can see how how clustering is first used to input data to determine clusters which are used as classes in classification later.

Possible applications of clustering are marketing: finding groups of customers with similar behavior given a large database of customer data; earthquake studies: clustering observed earthquake epicenters to identify dangerous zones; and others [2].

Possible applications of classifications are text categorization, medical diagnosis, protein function classification. [3]

Reference:

1. Lavine, B. K., Mirjankar, N. (2000). Clustering and classification of analytical data. Encyclopedia of analytical chemistry, 11.
2. Madhulatha, T. S. (2012). An overview on clustering methods. arXiv preprint arXiv:1205.1117.
3. Tsoumakas, G., Katakis, I. (2007). Multi-label classification: An overview. International Journal of Data Warehousing and Mining (IJDWM), 3(3), 1-13.