Purpose
To study classification and novelty detection using self-organising feature maps on a real application, using odour data collected with an electronic nose.

Presentation
The results from the exercise should be presented individually to one of the lab organizers. If a question or sub-task appears with a box then your respective answer must be presented. Students must be present at the next lab to answer questions on their work (the deadline for completion is the start of the next lab). If two students work together in a pair, then both students must be present and able to demonstrate the software that they have written and explain the answers.

Data and Matlab functions
Data and m-files can be downloaded from the course homepage (http://www.aass.oru.se/~tdt/ann). Note: Read each task carefully before starting to solve it.

Data collection
In this task you will collect data with an electronic nose (see Figure 1). The collection will be done in the AASS Laboratory. Your group will collect odour samples from five different liquids. Full instructions on how to operate the nose will be given by the lab assistant.

Task 1, SOM classifier
In this task you will use the SOM as a classifier for the different odours. The data from all groups will be provided on the course homepage after data collection. To complete this task, you should do the following:

1. Divide the collected data into separate data sets for training and testing.
2. Train and label a SOM in the same way as in Lab 5 using the training data.
3. Test the SOM on the training data, using the `som_bmu` function. Give the percentage of correctly classified vectors. Does the SOM classify the training data as expected?

________________________________________________________________________

4. Test the SOM on the test data. Give the percentage of correctly classified vectors. Does the SOM classify the data as expected?

________________________________________________________________________

5. Try to classify the unknown data with the SOM. Explain the results.

________________________________________________________________________
**Task 2, Novelty detection**

In this task the SOM will be used as a novelty detector. A SOM is trained using data from only one class. When testing the SOM all the data is used. The command:

```
[Bmus, Qerrors] = som_bmus(sm, data);
```

will produce two vectors where the elements in `Qerrors` show the Euclidean distance between each item in the data set and the best matching unit in the SOM. A shorter distance means less novelty compared to the SOM, meaning that a pattern is more closely related to the training data than the samples with larger distances.

1. Train and label one SOM using data representing only one of the classes.

2. Test the SOM with all the test data. Explain the result:

3. Does the SOM work as a novelty detector? Explain why/why not:

**Task3, Ensemble of SOMs (ESOM)**

In this task an ensemble of SOMs (ESOM) will be used as a classifier. This means that instead of using only one SOM, a collection of SOMs will be used – one for each class in the training data. By training each of the SOMs with data from one class (i.e., data from class 1 is used to train SOM 1, etc.) an ESOM is created.

To use the ESOM as a classifier, the data should presented to the ESOM, then the SOM with the shortest Euclidean distance to its best matching unit will be the “winner” (i.e., if SOM 2 has the shortest distance, then the sample belongs to class 2).

1. Train one SOM for each class.

2. Test the ESOM on the training data. Give the percentage of correctly classified vectors. Comment on the result:

3. Test the ESOM on the test data. Give the percentage of correctly classified vectors. Comment on the result:
4. Classify the unknown data. Compare the results with the result from task 2.

5. Compare the performance of the ESOM with that of the single SOM classifier in task 1. Your conclusions?