Goals
A frequently used kernel for learning from sequential data are sequence kernels. The aim of the exercise is to implement such a sequence kernel for various distance measures and apply this kernel to data.

Problem setting
The Dynamic Time Warping Kernel is defined by:

\[ k(x, y) = \exp(-\lambda D_{DTW}(x, y)) \]

where \( D_{DTW}(x, y) \) is the DTW-distance between to sequences \( x = \{x_i\} \) with \( i = 1 \ldots t_x \) and \( y = \{y_j\} \) with \( j = 1 \ldots t_y \). The DTW-distance is recursively defined by:

\[ \gamma(i, j) = d(x_i, y_j) + \min(\gamma(i - 1, j - 1), \gamma(i - 1, j), \gamma(i, j - 1)) \]

There are the following different DTW-distances \( d(x_i, y_j) \):

- **Edit distance**: binary distance between the data points \( x_i \) and \( y_j \), that is \( d(x_i, y_j) = [x_i \neq y_j] \).
- **Euclidean DTW distance**: (quadratic) euclidean distance between the data points \( x_i \) and \( y_j \), that is \( d(x_i, y_j) = (x_i - y_j)^2 \).
- **Derivative DTW-distance**: (quadratic) euclidean distance between the derivatives at the points \( x_i \) and \( y_j \), that is \( d(x_i, y_j) = (\Delta(x_i) - \Delta(y_j))^2 \), where the derivation can be calculated approximately by

\[ \Delta(x_i) = \frac{1}{2} \left( x_i - x_{i-1} + \frac{x_{i+1} - x_{i-1}}{2} \right) \]

Task 1
Write the following MATLAB-Function

```matlab
function D = dtw(x, y, dst)
```

which computes the DTW-distance between the sequences \( x \) and \( y \). Both sequences are given by vectors. The parameter \( \text{dst} \) controls, which distance is used: \( \text{dst}=1 \) stands for the edit distance, \( \text{dst}=2 \) stands for the euclidean distance, and all other parameters of \( \text{dst} \) stands for the derivative distance.
Task 2
Write the following MATLAB-Function

```
function K = dtw_kernel(X,Y,dst,lambda)
```

which computes the DTW-kernel. The matrix $X$ is a $n_X \times m_X$ matrix with $n_X$ sequences of the length $m_X$. The matrix $Y$ is a $n_Y \times m_Y$ matrix with $n_Y$ sequences of length $m_Y$. The parameter $dst$ controls, which distance is used and the parameter $\lambda$ is the kernel parameter. For the matrix $K$ the entry $K_{ij}$ is defined as $K_{ij} = \exp(-\lambda D_{DTW}(X_i, Y_j))$, where $X_i$ is the $i$’s row of matrix $X$ and $Y_i$ is the $j$’s row of matrix $Y$.

Task 3
Analyze the MATLAB functions `alpha = learnKernelRegERM(X,y,lambda,f_krn1,krnl_p)` and `y = classKernelRegERM(alpha,Xtr,f_krn1,krnl_p,X)`, which learns the parameters for a kernel model and classifies data for a previously learned model, respectively.

Task 4
Use the data set `laser_small.mat` and divide this data set in a training and test set using the function `split_train_test (split= 0.7, seed= 3, ...).` Train a model using the instruction `alpha = learnKernelRegERM(train,y_train,1,@dtw_kernel,0.0005)`.
Classify the test data and compute the accuracy for the previously trained model.