CLASSIFICATION WITH ARTIFICIAL NEURAL NETWORKS AND SUPPORT VECTOR MACHINES: APPLICATION TO OIL FLUORESCENCE SPECTRA

Khaled Mohamed Almhdi[1,3], Paolo Valigi[1], Vidas Gulbinas[2], Rainer Westphal[3] and Rainer Reuter[3]

1. Universita' di Perugia, Dipartimento di Ingegneria Elettronica e dell'Informazione, 06125 Perugia, Italy
2. Institute of Physics, 2053 Vilnius, Lithuania
3. Carl von Ossietzky Universität Oldenburg, Institut für Physik, 26111 Oldenburg, Germany

ABSTRACT

This paper reports on oil classification with fluorescence spectroscopy. The investigations are part of the development of a laser-based remote sensor (laser fluorosensor) to be used for the detection and classification of oil spills on water surfaces. The polychromator of the fluorosensor has six channels for measuring signals that represent the spectral fluorescence signature of the detected oil in the UV/VIS wavelength range following excitation at 355 nm wavelength. The investigation of the oil classification is based on the shape of the signature of the oil detected by these channels. The investigation uses three methods to examine crude oils, heavy refined oils, and sludge oils: the channels' relationships method (CRM); artificial neural networks (ANNs); and support vector machines (SVMs). This was done on a laboratory database of oil fluorescence spectra.

The database and the input fluorescence signature of the oils play a very important role in the efficiency of the classification method. If the input fluorescence of the oil does not fit into one of the classes already included in the database or if it is a completely new and previously not considered signature, then the training process for classification must always be redone. Generally, all three methods yield promising results and can be used for the detection and classification of oil spills on water surfaces. The channels' relationship method provides a meaningful classification of the available spectra, according to a rough oil type estimation. More specific substance information can be achieved with ANNs and SVMs. Both SVMs and ANNs prove to be efficient, fast, and reliable and have real-time capabilities. The SVM method is faster and more stable than ANN. Therefore, it is considered to be the most convenient method for classifying spectral information.

Keywords: Artificial Neural Networks, Multi Layer Perceptron, Support Vector Machines, oil fluorescence.

INTRODUCTION

Artificial neural networks (ANN) were involved in many applications to solve real world problems. In commercial purposes, ANNs can be applied to predict the profit, market movements, and price levels based on the market's historical dataset. In medical applications, doctors can evaluate the situation of many patients depending on the historical dataset of other patients who had the same illness. In industry, engineers can apply ANNs to solve many engineering problems such as classifications, prediction, pattern recognition, and non-linear problems where the issues are very difficult or might be impossible to solve through normal mathematical processes. ANNs were applied to predict slant path rain attenuation (1), to predict the rain attenuation on an Earth-space path, to predict the water quality index (WQI), and to signal predictions in a nuclear power plant (2). They were also used in face recognition (3). In medical applications, ANNs were utilised in detecting brain disease (4) and DNA ploidy, as well as cell cycle distribution of breast cancer aspirate cells that are measured by image cytometry and analysed by ANNs for their prognostic significance (5).

Support vector machines (SVM) are modern and effective tools that have already been examined to solve difficulties such as classification problems and pattern recognition. In SVMs there is no need to select features from several applications. SVMs were applied to medical binary classifica-

tion problems (6), to recognise radar emitter signals (7), and to visual speech recognition (8), and in many other cases.

This paper reports on oil classification with fluorescence spectroscopy. The objective is to classify the oil fluorescence spectra based on a laboratory dataset of fluorescence spectra of several oil classes (sludge, crude and heavy oil). The classification was carried out using the following three methods: channel relationship method (CRM), artificial neural networks (ANNs), and support vector machines (SVMs).

METHODS

Within the EU-funded project FLUOSENSE a laser fluorosensor has been developed for the detection of oil films on water surfaces via fluorescence excitation over distances of 50 to 100 m (9). The instrument consists of a UV-emitting (355 nm wavelength) pulse laser for target illumination, a telescope for the efficient collection of light from the illuminated area, and a polychromatic grating spectrometer equipped with photomultipliers for a sensitive detection of spectral information in six UV/VIS channels at 366-395 nm (channel 1), 386-425 nm (channel 2), 442-486 nm (channel 3), 492-552 nm (ch annel4), 574-644 nm (channel 5) and 630-700 nm (channel 6) (Figure 1).

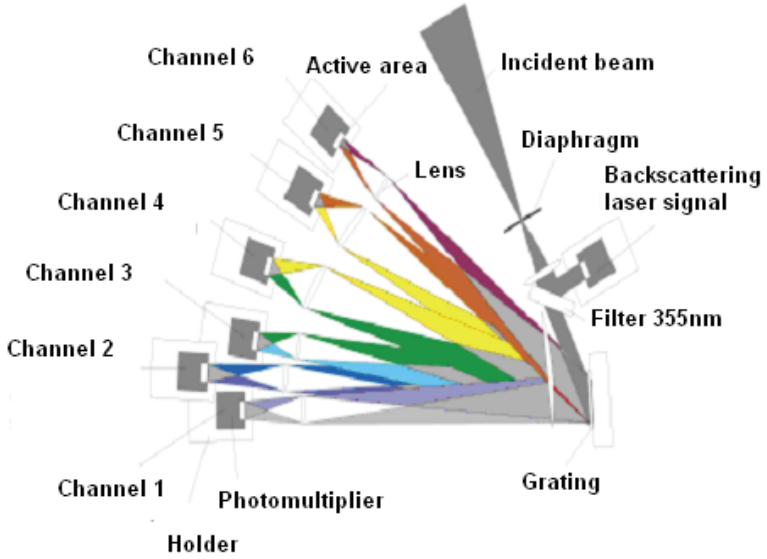


Figure 1: Optical design of the polychromator.

For examining the oil type classification capability of the spectrograph, a dataset of 117 spectra of 16 sludge, 80 crude and 21 heavy refined oils were measured with a Perkin Elmer LS50 laboratory spectrofluorometer. The measuring procedure used for this dataset is the same as documented in an earlier data catalogue (10) but includes also the 355 nm excitation wavelength which was not considered in (10). Each substance sample is analysed with 355 nm excitation wavelength and 386-700 nm emission bandwidth. The geometrical orientation of the sample cuvette with respect to the excitation and emission ray path differs from the conventional 90 degree configuration. Because of the high absorbance of samples, the cuvette is oriented such that fluorescence is excited and registered through a single side of the cuvette, with a cuvette holder termed *front surface assembly* by some manufacturers (10). This configuration corresponds well to the geometry of laser fluorosensor measurements of oil films on water, and therefore the dataset can be used as a basis for fluorosensor data interpretation.

The absolute fluorescence efficiency of the oils is of minor relevance since the signal intensities measured with the laser fluorosensor would depend on other factors as well, e.g. on the target distance and the oil film thickness. Therefore, spectral intensities are discarded by normalising the integral over the entire emission spectrum to the value of 1, and these normalised spectral signatures are used for classification. In the next step, the normalised spectra are integrated in the limits of the detection channels 2 to 6, to simulate data which are measured by the laser fluorosensor in

the presence of these oils (Figure 2). Channel 1 is not further considered here since in practical fluorosensor applications over water this signal includes the water Raman scattering which is useful for oil film thickness measurements (11,12) and due to this underlying signals less specific for oil type classification.

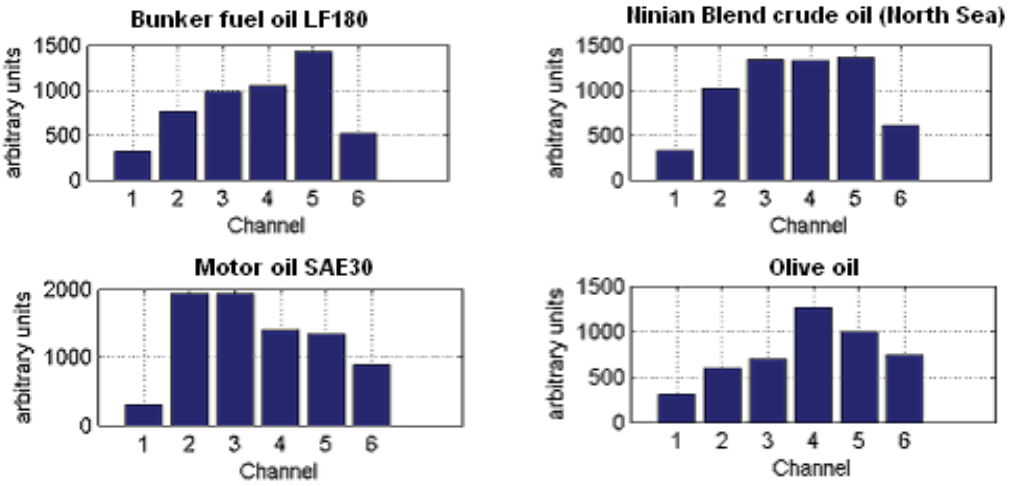


Figure 2: Examples of oil fluorescence spectra (355 nm excitation) spectrally grouped according to the detection channel wavebands of the laser fluorosensor.

The classification is carried out using three methods: artificial neural networks (ANNs), support vector machines (SVMs), and channel relationship method (CRM). The normalised fluorescence signatures detected by the channels of the polychromator are considered to be the inputs to the classification method to determine their output (oil class).

Channel Relationship Method (CRM)

The channel relationship method (CRM) is a method relating the relative intensity of the normalised oil fluorescence detected by the polychromator channels. Some fluorescence signals have a strong appearance at a certain wavelength and a weak appearance in other wavelength ranges. Fluorescence signals which have shapes similar to each other form their own group in a diagram. We do expect that such oils where the fluorescence signatures are close to each other can be differentiated from other oil classes. In this way it is expected that several oil classes can be identified, and only those oils that have signatures which are close to more than one class lead to difficulties in evaluating the class of such oils.

Artificial Neural Networks (ANN)

Artificial neural networks are powerful tools that can learn to solve problems in a way similar to the human brain. ANNs gather knowledge by detecting the patterns and relationships in data and learn (or: are trained) through experience, not from programming (13). Figure 3 shows the structure of the ANN. It is a combination of many single neurons. The ANN might consist of several thousand artificial neurons, and the output of one neuron becomes an input to another neuron. Figure 4 shows the neuron model of the artificial neural network where the output of such a neuron is given by:

$$Z = f\left(\sum_{0}^{4} \omega_i X_i\right) \tag{1}$$

where ω_i is defined as the weight and X_i is the input to the networks.

Many transfer functions could define $f(x)$ (Figure 5). The transfer function used in this work is the tanh function which is among the most popular functions to the neural network design due to its mathematical properties such as monotonicity, continuity, and differentiability, which are important to the training process (14).

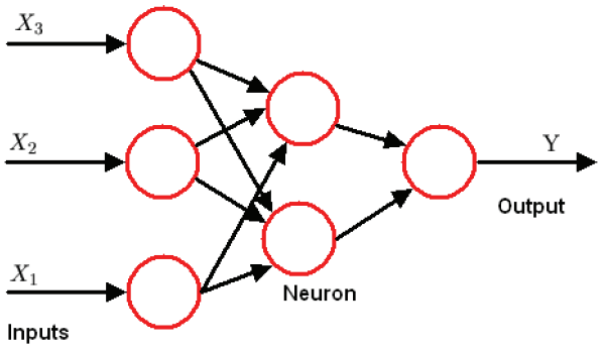


Figure 3: Structure of an Artificial Neural Network.

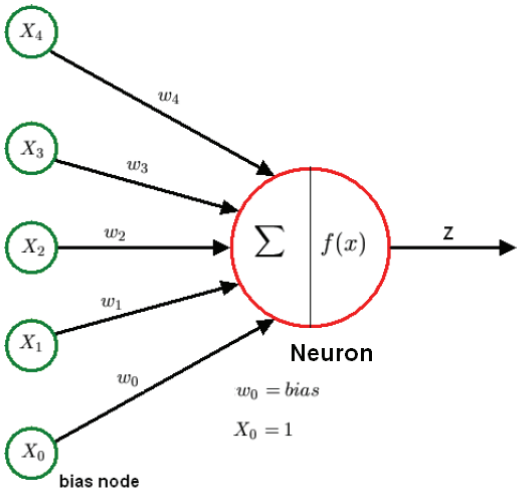


Figure 4: Neuron Model of an Artificial Neural Network.

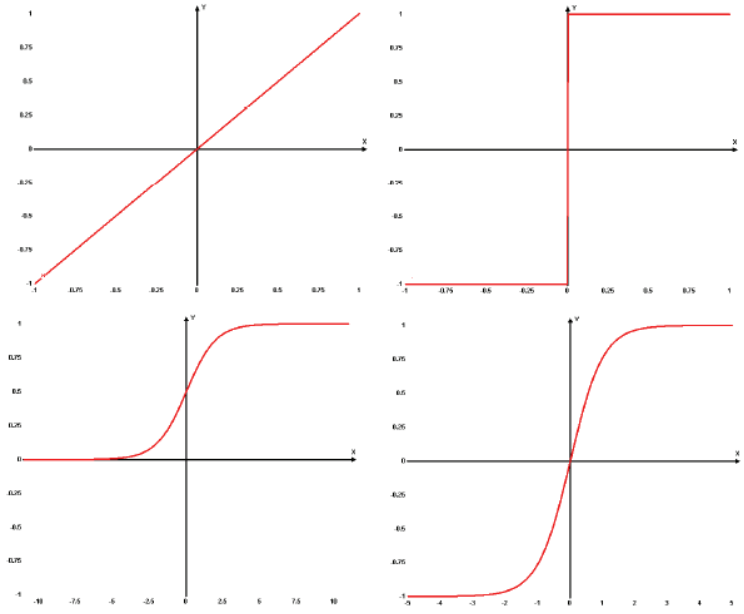


Figure 5: Transfer functions to neural network design.

There are several types of ANNs according to their structure and learning algorithms. According to their structure, ANNs can be classified as feed forward networks and recurrent networks (15). In a feed forward network, the neurons are generally grouped into layers. Signals flow from the input layer through the output layer via unidirectional connections, the neurons being connected from one layer to the next, but not within the same layer (15). In recurrent networks, the output of some

neurons is fed back to the same neurons or to neurons in a preceding layer (15). Concerning the learning process, the ANN can be classified according to supervised learning, unsupervised learning or reinforcement. In the supervised model, the ANN requires the output in order to adjust its weight. In the unsupervised model, the ANN does not require the output, the ANN adapts purely in response to its input. These networks learn and build their structure based on the input. The reinforcement learning algorithm employs a critic to evaluate the quality of the neural network output corresponding to a given input (15).

In this paper, the ANN shall estimate the unknown oil class in the laboratory dataset of the different oil classes. Multilayer perceptron (MLP) is a well-known type of ANN, which is usually used in classification problems. In MLP, the neurons are grouped in many layers (Figure 6). In this work, MLP with supervised learning has been used. In this approach, during the training process of the network, the network compares its actual results $y(t)$ with the desired output $d(t)$ and then computes the error (Eq. 2). Through the backpropagation algorithm the error is presented many times to the input of the forward activation place, and the process continues until the actual outputs get closer to the desired output. The MLP that is used in this work contains three layers: input layer, one hidden layer, and output layer. Eq. 2 represents the error calculated for the input vector that is presented to the feed forward network and Eq. 3 represents the error calculated for all input vectors that are presented to the feed forward network.

$$\varepsilon(t) = \frac{1}{2}\sum_{i}(d_i(t) - y_i(t))^2 \tag{2}$$

$$\varepsilon_t(t) = \frac{1}{2}\sum_{v=1}^{v}\sum_{i}(d_i(t) - y_i(t))^2 \tag{3}$$

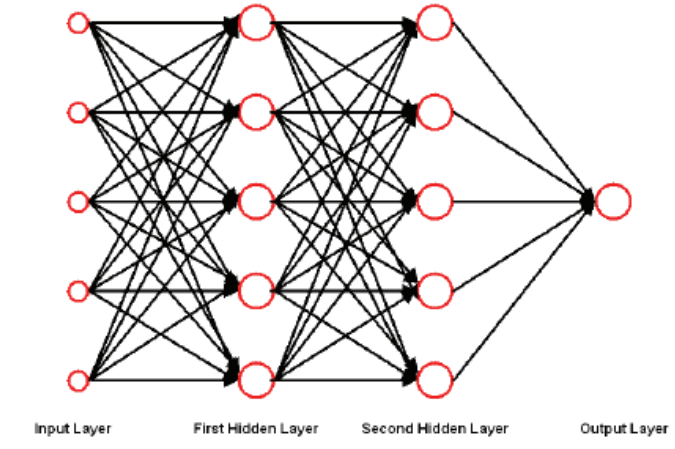


Figure 6: Multilayer Perceptron (MLP).

The NeuroSolutions software (NeuroDimension Inc., Gainsville, Florida, USA) is used in this work to construct the required MLP (16). The constructed network consists of an input layer, a nonlinear hidden layer, and an output layer. The hidden layer and the output layer apply the *tanh* transfer function.

The MLP network has been trained based on 111 oil fluorescence signatures of crude, heavy and sludge oils. The cross validation process of the network uses a dataset of nine oil fluorescence spectra from the same oil classes. Cross validation is an optional method, it is a highly recommended method for stopping network training. This method monitors the error on an independent set of data and stops training when this error begins to increase. This is considered to be the point of best generalisation (16). The testing process is defined as data used to evaluate the performance after the training is complete (16). The trained network then was tested based on three oil fluorescence signatures that were not used in the training set from the same three oil classes used in the training in order to ensure that the trained network understood the problem. Then, three un-

known oil fluorescence spectra that were also not used in the training and testing sets from the same three oil classes used in the training were fed into the trained network for classification. The trained MLP was retrained using the same dataset and then tested. The three unseen spectra were re-classified and the results compared in order to estimate the quality of the result and to investigate the stability of the trained MLP. The number of nodes of the network is the number of exemplars (rows) of the training set equal to 111 and 1000 epochs. Note that the number of nodes is configured automatically by the NeuralBuilder (16).

Support Vector Machines (SVM)

SVMs are powerful tools for classification that can be considered as an alternative to the mulitlayer perceptron. SVMs were first introduced in 1992 (17). A good explanation for the SVMs theory and its applications is found in (18). The basic idea is to find the linear classifier called the hyperplane. Figure 7 shows many linear classifiers separated in two classes (red and green). There is an ideal separating classifier (black line) called hyperplane (maximum margin linear classifier) which can increase the space between it and the nearest dataset points of different classes as much as possible. For this, SVMs are regarded to be margin classifiers. In cases where the classes can be separated easily, this kind of support vector machines is called linear support vector machine (LSVM).

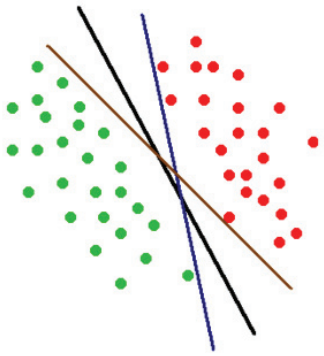


Figure 7: The best linear classifier (hyperplane).

For separable classes (Figure 8), a SVM classifier computes a decision function with a maximal margin *M* with respect to the two classes (blue and red classes). There are two planes touching the boundary of dataset, $w\,x+b=+1$ and $w\,x+b=-1$. w is a vector perpendicular on the plane $w\,x+b=+1$. We consider x_2 to be any point on the plane $w\,x+b=+1$ and x_1 to be the closest point to x_2 on the plane $w\,x+b=-1$. The line from x_2 to x_1 is perpendicular to the planes. The maximum margin of the best classifier can be expressed as $M = 2/\sqrt{w \cdot w}$.

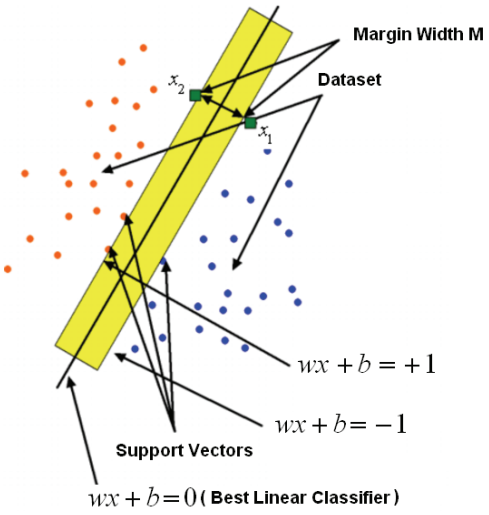


Figure 8: Classification of a linearly separable dataset.

We consider a dataset: $(x_1, y_1), \ldots\ldots\ldots, (x_n, y_n), \quad \forall_i \in (1, \ldots\ldots, n)$. The decision boundaries can be found by solving the following constrained optimising problem:

Minimise $\frac{1}{2}\|w\|^2$ subject to: $y_i(w \cdot x_i + b) - 1 \geq 0, \quad \forall_i$ (4)

The Lagrange Function Formulation of this optimising problem is given by:

$$L(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_i^n \alpha_i (y_i(w \cdot x_i + b) - 1), \quad \alpha_i \geq 0 \quad \forall_i \tag{5}$$

by setting the derivative of the Lagrange Function to zero:

$$\frac{\partial}{\partial b} L(w, b, \alpha) = 0, \qquad \frac{\partial}{\partial w} L(w, b, \alpha) = 0$$

This leads to the conditions: $\sum_{i=1}^n \alpha_i y_i = 0$ and $w = \sum_{i=1}^n \alpha_i y_i x_i$

substituting into $L(w, b, \alpha)$, the optimisation problem can be expressed as:

$$\text{Max. } W(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle \tag{6}$$

subject to: $\alpha_i \geq 0, \qquad i = 1, \ldots\ldots\ldots, n \qquad \sum_{i=1}^n \alpha_i y_i = 0$

x_i with non zero value of α_i are called support vectors:

$$y_i \left[\langle w, x_i \rangle + b\right] = 1 \quad \Rightarrow \quad \alpha_i \triangleright 0 \tag{7}$$

In this case α_i is a margin. $\Rightarrow$ x_i are called support vectors.

Or $$y_i \left[\langle w, x_i \rangle + b\right] \triangleright 1 \quad \Rightarrow \quad \alpha_i = 0 \tag{8}$$

$\Rightarrow$ x_i are not considered to be support vectors.

In the case of a non-linearly separable dataset (Figure 9), the slack variable ζ_i is introduced leading to a soft margin classifier:

Minimise $\frac{1}{2}\|w\|^2 + C\sum_{i=1}^n \zeta_i$

subject to: $y_i(w \cdot x_i + b) \geq 1 - \zeta_i \qquad \forall_i$ (9)

The parameter C describes the trade-off between the maximal margin and the correct classification (18). This idea leads to the following dual problem

$$\text{Max. } W(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle \tag{10}$$

subject to: $C \geq \alpha_i \geq 0, \qquad i = 1, \ldots\ldots\ldots, n \qquad \sum_{i=1}^n \alpha_i y_i = 0$

One can note that it is similar to an optimisation problem in case of linearly separable data with the exception of α_i being limited to the upper bound C.

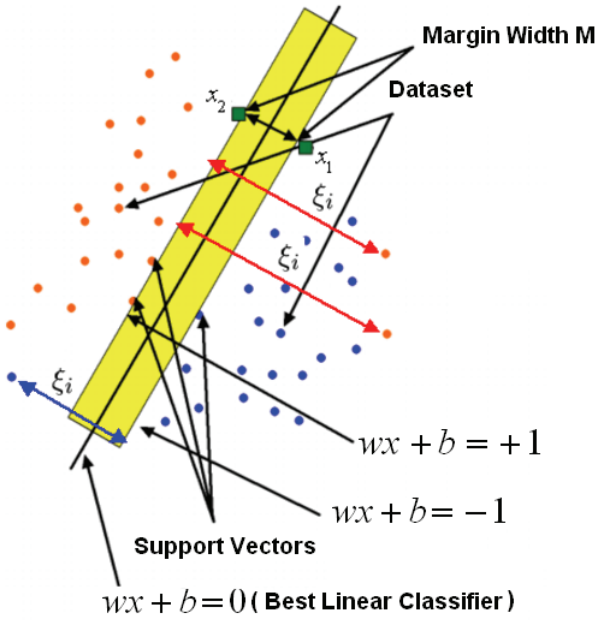


Figure 9: Classification of a non-linearly separable dataset.

To solve non-linear classification problems, linear support vector machines are applied to high dimensional spaces (Figure 10). Transforming data into a high dimensional space can transform complex problems (with complex decision surfaces) into simpler problems that can be solved with linear classifiers (16). This means transforming from $\langle x_i, x_j \rangle$ to $\langle \Phi(x_i), \Phi(x_j) \rangle$ (Kernel trick $K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$ leading to the following dual problem:

$$\text{Max. } W(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \tag{11}$$

The common Kernel functions are polynomial with degree d, the Radial Base Function with width σ and the Sigmoid with parameter k

$$K(x_i, x_j) = \left(\langle x_i, x_j \rangle + C \right)^d \tag{12}$$

$$K(x_i, x_j) = \exp\left(-\frac{1}{2\sigma^2} \| x_i - x_j \|^2 \right) \tag{13}$$

$$K(x_i, x_j) = \tanh\left(k \langle x_i, x_j \rangle + \Theta \right) \tag{14}$$

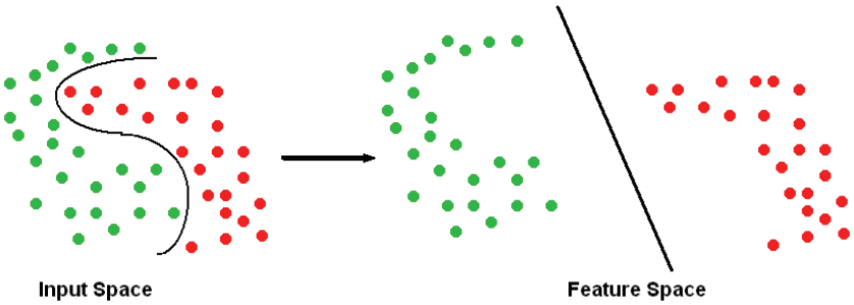


Figure 10: Transforming datasets from the input space to the high dimensional space.

As in the case of ANNs, a NeuroSolutions programme was used in this work (16). In this programme, transforming the data from input space to high dimensional space is done using a Radial Basis Function (RBF) network that places a Gaussian distribution at each data sample (16). Thus, the feature space becomes as large as the number of samples (16). The SVM was divided into two parts to implement the RBF dimensionality expansion and a large margin classifier. As in the case of MLP, SVMs use the concept of backpropagation training to train the linear combination of Gaussians. SVMs are motivated by the concept of training and use only those inputs that are near the decision surface since they provide most information about the classification (16). More details on the components of the constructed SVM are available in (16).

The training, testing, and production processes of SVM were carried out based on the same dataset that was used in MLP to ensure an exact comparison in the quality of the results between the MLP and SVM. As in MLP, the number of nodes of an SVM network is the number of exemplars (rows) of the training set equal to 111 and 1000 Epochs. The number of nodes is configured automatically by the NeuralBuilder (16).

Classification of noisy oil spectra

Meeting the practical situation of noisy oil fluorescence data requires investigating the noise's effect on the oil type classification. This was achieved by two approaches. The first is to introduce noisy oil fluorescence data to be classified by the MLP and SVM that were trained on the ideal oil spectra (non-noisy spectra) measured in the laboratory in order to see the trained MLP's and SVM's ability to classify them in the presence of noise. The second approach is to expand the dataset to include both non-noisy and noisy spectral data. The results achieved in both approaches are compared.

Use of the noise-free dataset

The same three new non-noisy input oil spectra that were classified with the trained MLP and SVM were presented once again to the same trained MLP and SVM, but as noisy spectra. For this, 5% and 10% random Gaussian noises were added to these new input spectra and then presented to the trained MLP and SVM.

Training with noisy spectra

To ensure the best performance of the trained MLP and SVM the original dataset was expanded to include, in addition to the ideal spectra (without noise), the same spectra after the 10% random Gaussian noise had been added to them. I.e., the original dataset was doubled to include also noisy spectra with a random Gaussian noise of 10%. Then, the same three new noisy spectra were presented to the trained MLP and SVM.

RESULTS AND DISCUSSION

Oil classification and identification using the CRM

Figure 11 shows the results of the oil classification using the CRM.

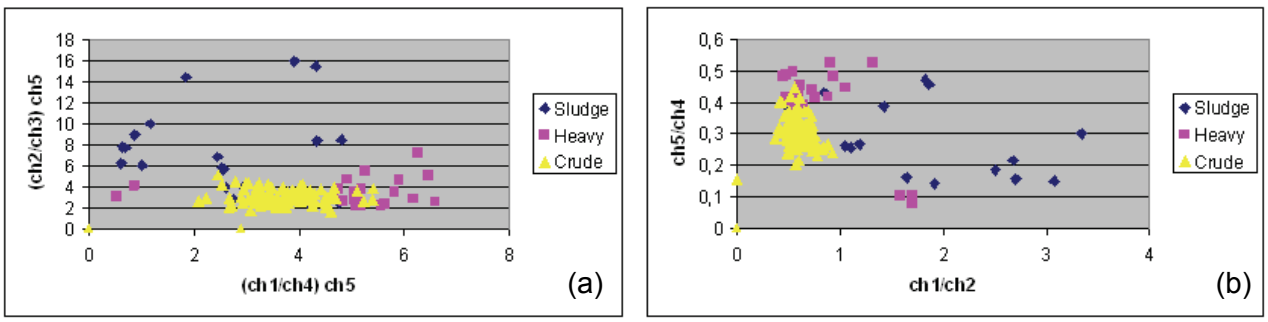


Figure 11: Oil classification using the CRM.

Most of the oils within the crude oil class have similar fluorescence spectral signatures. Therefore, the oils within this class form their own group when the CRM is applied. The same role is applied to heavy oil. Oils within the sludge oil classes are randomly distributed since the oils within this class have different oil fluorescence spectra. It is also shown that the spectra of some oils from the heavy and crude oils resemble the sludge oil spectra, which makes them mixed within the two classification groups. So, if the detected oil fluorescence appears in a yellow area crude oil will be mostly expected, and if it appears in a red area heavy oil will be expected, otherwise, it will be expected to be sludge oil. This classification is based on the assumption that there are only three oil classes. Other types of oil classes such as bilge oil and vegetable oil could be considered to increase the ability of this method to classify oil fluorescence spectra.

Oil classification and identification using MLP

Table 1 shows the testing results of the trained network. The output of the tested MLP shows that the closest value to 1 is for sludge oil as the first input, while heavy oil is the second and crude oil is the third input. This result meets the desired output performance of the trained MLP. Table 2 shows the identification of unknown oil spectra by the trained network. There were three new fluorescence spectral signatures of oil, which the trained MLP was unaware of. The result shows that the first oil is sludge, the second is crude, and the third is heavy oil. These identifications were made depending on which class had the closest value to 1. On the basis of the database the trained MLP classified the three new fluorescence spectral signatures of oil correctly.

Tables 3 and 4 show the results of the retrained network. From the results one can see that the classification is still correct, but the accuracy of the classification is not the same although both networks were trained on the basis of the same dataset. This means that one should train the network many times to evaluate the best performance of the trained network.

Table 1: Testing results of the trained MLP.

Ideal Output			Calculated Output		
Sludge	**Heavy**	**Crude**	**Sludge**	**Heavy**	**Crude**
1.000	0.000	0.000	0.863	0.279	0.001
0.000	1.000	0.000	0.022	0.844	0.228
0.000	0.000	1.000	0.052	-0.017	0.949

Table 2: The identification of unknown oil spectra by the trained MLP. The used spectra were from sludge, crude and heavy oil.

Calculated Output		
Sludge	**Heavy**	**Crude**
0.869	0.226	-0.025
0.083	-0.006	0.946
0.036	0.7019	0.329

The following results are those of the re-trained MLP.

Table 3: Testing results of the re-trained MLP.

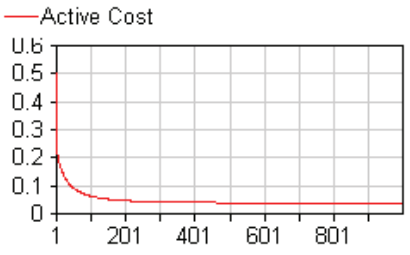
Ideal Output			Calculated Output		
Sludge	**Heavy**	**Crude**	**Sludge**	**Heavy**	**Crude**
1.000	0.000	0.000	0.782	0.205	-0.011
0.000	1.000	0.000	-0.012	0.751	0.232
0.000	0.000	1.000	-0.002	-0.045	0.992

Table 4: Identification of unknown oil spectra by the re-trained MLP. Same spectra as in Table 2.

Calculated Output		
Sludge	**Heavy**	**Crude**
0.858	0.270	-0.030
0.020	-0.043	0.978
0.012	0.634	0.338

Oil classification and identification using a SVM

Figure 12a shows the active cost of the SVM. One can see that the cost curve of the SVM approaches to zero, which means that SVM has understood the problem. Table 5 shows the testing results of the trained SVM. The output of the tested SVM shows the closest value to 1 for sludge oil at the first input, for heavy oil at the second and for crude oil at the third input. This result indicates that the trained SVM meets the desired requirements. Table 6 shows the results of the identification by SVMs. These results were achieved from the first time training. The re-training process has no effect on the result. This means the performance of the SVM is achieved from the first time training process which saves the user's time. The identification of an unknown oil fluorescence spectrum is 100% correct. These identifications were made based on which class had the closest value to 1, as it was done with the MLP. When comparing these methods one can see that both MLP and SVMs correctly identified the unknown spectra.



(a)

MSE	0.007951633209
NMSE	0.012936453271
r	0.999400724529
% Error	1.489958023330
AIC	1461.384049074099
MDL	1815.794384629610

(b)

Figure 12: (a) The active cost curve approaches zero which means that classification of the dataset was carried out correctly. (b) the performance, showing the mean squared error (MSE), the normalized mean squared error (NMSE), percent error (% error), Akaike's information criterion (AIC), and Rissanen's minimum description length (MDL) criterion.

Table 5: Testing results of SVM.

Ideal Output			Calculated Output		
Sludge	**Heavy**	**Crude**	**Sludge**	**Heavy**	**Crude**
1.000	0.000	0.000	0.902	0.065	0.099
0.000	1.000	0.000	0.124	0.817	0.163
0.000	0.000	1.000	0.046	0.060	0.946

Table 6: Identification of unknown oil spectra by a SVM. Same spectra as in Table 2.

Calculated Output		
Sludge	**Heavy**	**Crude**
0.844	0.103	0.133
0.052	0.071	0.937
0.047	1.097	-0.054

CLASSIFICATION OF NOISY OIL SPECTRA

Use of the noise-free dataset

Tables 7-10 show the identification of new noisy oil spectra that were presented to the trained MLP and SVM. These results show that the trained MLP and SVM identified the noisy inputs, although the MLP and SVM were trained with the noise-free dataset. Again the identification is based on the output whose value is closest to 1. Thus, the first input sample would be identified as sludge oil, the second one as crude oil, and the third one as heavy oil. This identification is correct according to the used types of oils.

Table 7: Identification of unknown noisy spectra by the trained MLP. Same spectra as in Table 2 after adding 5% random Gaussian noise.

Calculated Output		
Sludge	**Heavy**	**Crude**
0.933	0.182	0.120
0.028	-0.047	0.997
-0.041	0.794	0.194

Table 8: Identification of unknown noisy spectra by the trained MLP. Same spectra as in Table 2 after adding 10% random Gaussian noise.

Calculated Output		
Sludge	**Heavy**	**Crude**
0.891	0.266	0.069
0.044	-0.047	0.995
-0.036	0.604	0.337

Table 9: Identification of unknown noisy spectra by the trained SVM. Same spectra as in Table 2 after adding 5% random Gaussian noise.

Calculated Output		
Sludge	**Heavy**	**Crude**
0.681	0.186	0.266
0.287	0.276	0.627
0.282	0.508	0.438

Table 10: Identification of unknown noisy spectra by the trained SVM. Same spectra as in Table 2 after adding 10% random Gaussian noise.

Calculated Output		
Sludge	**Heavy**	**Crude**
0.651	0.202	0.291
0.187	0.184	0.760
0.239	0.665	0.302

Training with noisy spectra

Tables 11-14 show the identification of the new noisy oil spectra presented to the MLP and SVM that were trained with the enlarged dataset. These results show that the trained MLP and SVM could identify the noisy inputs correctly. As seen in the comparison between the results achieved

from the MLP and SVM that had been trained based only on the non-noisy spectra and the results achieved from the MLP and SVM that had been trained based on the dataset that contained non-noisy and noisy spectra, there is not much difference in the quality of the results.

Table 11: Identification of unknown noisy spectra by the trained MLP. Same spectra as in Table 2 after adding 5% random Gaussian noise.

Calculated Output		
Sludge	**Heavy**	**Crude**
0.954	0.007	-0.054
-0.012	-0.038	0.877
-0.027	0.845	0.470

Table 12: Identification of unknown noisy spectra by the MLP. Same spectra as in Table 2 after adding 10% random Gaussian noise.

Calculated Output		
Sludge	**Heavy**	**Crude**
0.953	0.001	-0.054
0.000	-0.043	0.894
-0.028	0.724	0.602

Table 13: Identification of unknown noisy spectra by the trained SVM. Same spectra as in Table 2 after adding 5% random Gaussian noise.

Calculated Output		
Sludge	Heavy	Crude
0.874	0.127	0.092
0.160	0.209	0.774
0.389	0.479	0.419

Table 14: Identification of unknown noisy spectra by the trained SVM. Same spectra as in Table 2 after adding 10% random Gaussian noise.

Calculated Output		
Sludge	**Heavy**	**Crude**
0.877	0.125	0.090
0.057	0.089	0.924
0.324	0.477	0.446

CONCLUSIONS

The CRM provides a simple but still meaningful classification of the available spectra according to the rough oil type. From the results obtained with all investigated methods one can conclude that the SVMs are the most convenient method for oil spectral signatures classification and identification due to the fact that they are faster and more stable than the ANNs. Re-trained ANNs lead to changes of the accuracy although the general result is still correct. Meanwhile, a re-training process in SVMs does not lead to any change in the final result of oil classification and identification. This result gives the SVMs the advantage of being more stable than ANNs. It also saves the user time because by using the SVMs one can get the best result from the first time training. Meanwhile,

when using ANNs, it is necessary to train the network many times in order to estimate the best performance. Meeting the practical situation of noisy oil fluorescence, both ANNs and SVMs proved to be efficient to classify noisy oil fluorescence of 5% and 10% Gaussian noise. To ensure the best performance of the trained MLP and SVM the training dataset should include noise-free and noisy fluorescence spectra.

The database and the input fluorescence signature of the oils play a very important role in the efficiency of the classification method. If the input fluorescence of the oil does not fit into one of the classes already included in the database, or if it is a new and previously not considered signature, then the classification method must always be redone. Both ANNs and SVMs have real time capabilities, however, they can be trained offline and then applied in real-time applications. As a result, we conclude that the SVM classification method is the most convenient method to classify oil spectra. Since the dataset plays an important role in the accuracy of the classification method, it would be beneficial to widen the dataset to include other oils such as bilge oils, vegetable oils and other oil-like chemicals which are transported at sea. This work is an offline task, therefore, it can be integrated with the real sensor.

There are many programmes of ANNs and SVMs available, therefore, it is not imperative to restrict the programme used in this work. One might use another version and compare the quality of the result.

ACKNOWLEDGEMENTS

The classification and identification of the oil fluorescence spectra were carried out using the NeuroSolution programme (www.nd.com).

REFERENCES

1 Hongwei Yang, Chen He, Hongwen Zhu & Wentao Song, 2000. Prediction of slant path rain attenuation based on artificial neural network. In: Circuits and Systems 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on Circuits and Systems, 1, 152-155, DOI 10.1109/ISCAS.2000.857049

2 Kim WJ, S H Chang & B H Lee, 1993. Application of neural networks to signal prediction in nuclear power plant. IEEE Transactions on Nuclear Science, 40(5): 1337-1341 DOI 10.1109/23.234547

3 Ming Zhang & J Fulcher, 1996. Face recognition using artificial neural network group-based adaptive tolerance (GAT) trees. IEEE Transactions on Neural Networks, 7(3): 555-567 DOI 10.1109/72.501715

4 Jervis B W, M R Saatchi, A Lacey, T Roberts, E M Allen, N R Hudson, S Oke & M Grimsley, 1994. Artificial neural network and spectrum analysis methods for detecting brain diseases from the CNV response in the electroencephalogram. IEEE Proceedings on Science, Measurement and Technology, 141(6): 432-440

5 Naguib R N G, H A M Sakim, M S Lakshmi, V Wadehra, T W J Lennard, J Bhatavdekar & G V Sherbet, 1999. DNA ploidy and cell cycle distribution of breast cancer aspirate cells measured by image cytometry and analyzed by artificial neural networks for their prognostic significance. IEEE Transactions on Information Technology in Biomedicine, 3(1): 61-69 DOI 10.1109/4233.748976

6 Yuchun Tang, Bo Jin, Yi Sun & Yan-Qing Zhang, 2004. Granular support vector machines for medical binary classification problems. 2004. In: Proceedings of the 2004 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB '04), 73-78

7 Gexiang Zhang, Weidong Jin & Laizhao Hu, 2004. Radar emitter signal recognition based on support vector machines. In: Control, Automation, Robotics and Vision Conference (ICARCV 2004), 2: 826-831. DOI 10.1109/ICARCV.2004.1468946

8 Gordan M, C Kotropoulos & I Pitas, 2002. Application of support vector machines classifiers to visual speech recognition. In: Proceedings of the International Conference on Image Processing, 3: III-129 - III-132. DOI 10.1109/ICIP.2002.1038921

9 Karpicz R, A Dementjev, Z Kuprionis, S Pakalnis, R Westphal, R Reuter & V Gulbinas, 2005. Laser fluorosensor for oil spot detection. Lithuanian Journal of Physics, 45(3): 213-218

10 Catalogue of Optical Spectra of Oils, 2005 (University of Oldenburg, Institute of Physics, Section Marine Physics), http://las.physik.uni-oldenburg.de/data/spectra/index.htm

11 Hoge F E & R N Swift, 1980. Oil film thickness measurement using airborne laser-induced water Raman backscatter. Applied Optics, 19: 3269-3281

12 Hengstermann T & R Reuter, 1990. Lidar fluorosensing of mineral oil spills on the sea surface. Applied Optics, 29: 3218-3227

13 Agatonovic-Kustrin S & R Beresford, 2000. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. Journal of Pharmaceutical and Biomedical Analysis, 22(5): 717-727

14 Kevin L. Priddy and Paul E. Keller, 2005. Artificial Neural Networks: An introduction, SPIE - The International Society for Optical Engineering, Bellingham, Washington 98227-0010 USA, 2-5, ISBN: 0-8194-5987-9

15 Duc Truong Pham & Liu Xing, 1999. Neural networks for Identification, Prediction and control, Intelligent Systems Laboratory,Systems Division, Cardiff School of Engineering, University of Wales,Cardiff.CF2 3TE, UK: 1-3

16 NeuroDimension, Inc. (Gainsville, Florida, USA) http://www.neurosolutions.com/downloads/documentation.html

17 Boser B E et al., A, 1992. Training Algorithm for Optimal Margin Classifiers, Proceedings of the Fifth Annual Workshop on Computational Learning Theory 5 144-152, Pittsburgh.

18 Lipo Wang, 2005. Support Vector Machines: Theory and Applications (Springer-Verlag, Berlin Heidelberg) 136-141, SPIN: 10984697