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

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

1. Universita' di Perugina, Dipartimento di Ingegneria Elettronica e dell'Informazione, 06125 Perugia, Italy
2. Institute of Physics, 2053 Vilnius, Lithuania
3. 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 (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 fluorescence spectral signature of the detected oil in the UV/VIS wavelength range following excitation at 355 nm. 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 (ANN); and support vector machines (SVM). This was done based 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 classification method must always be redone. Generally, all three methods look promising 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 could 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.

INTRODUCTION

Artificial neural networks (ANN) have been involved in many applications to solve real world problems. In commercial purposes ANNs can be applied to predict the profit, market movements, and price level based on the market's historical dataset. In medical applications, doctors can evaluate the case of many patients depending on the historical dataset of other patients who had the same case. 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 have been applied to predict slant path rain attenuation (1), to predict rain attenuation on an Earth-space path, to predict water quality index (WQI), and to signal predictions in a nuclear power plant (2). They have also been used in face recognition (3). In medical applications, ANNs have been utilized 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. SVMs can be used to solve more complex problems when compared to ANNs. In SVMs there is no need to select features from several applications, and SVMs have demonstrated that they are more accurate and stable than ANNs, which will be proven for fluorescence spectra classification later in this paper.

SVMs have been applied to medical binary classification problems (6), to recognize radar emitter signals (7), to detect complicated attacks (8), and to visual speech recognition (9), and many other applications.

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 (10). The instrument consists of a UV-emitting (355 nm wavelength) pulse laser for target illumination, a telescope for 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 (1), 386-425 nm (2), 442-486 nm (3), 492-552 nm (4), 574-644 nm (5) and 630-700 nm (6) (Figure 1).

For examining the oil type classification capability of the spectrograph, a dataset of 111 spectra of sludge, crude and heavy refined oils 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 (11) but includes also the 355 nm excitation wavelength which was not considered in (11). The absolute fluorescence efficiency of the oils is of minar 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 spectra in the database to unity of their integral over the entire emission spectrum, 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 since in practical fluorosensor applications over water this signal includes the water Raman scattering which is seful for oil film thickness measurements (12,13) and due to this underlying signal less specific for oil type classification.

The classification is carried out using three methods: artificial neural networks (ANNs), support vector machines (SVMs) and channel relationship method (CRM). The normalized 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).

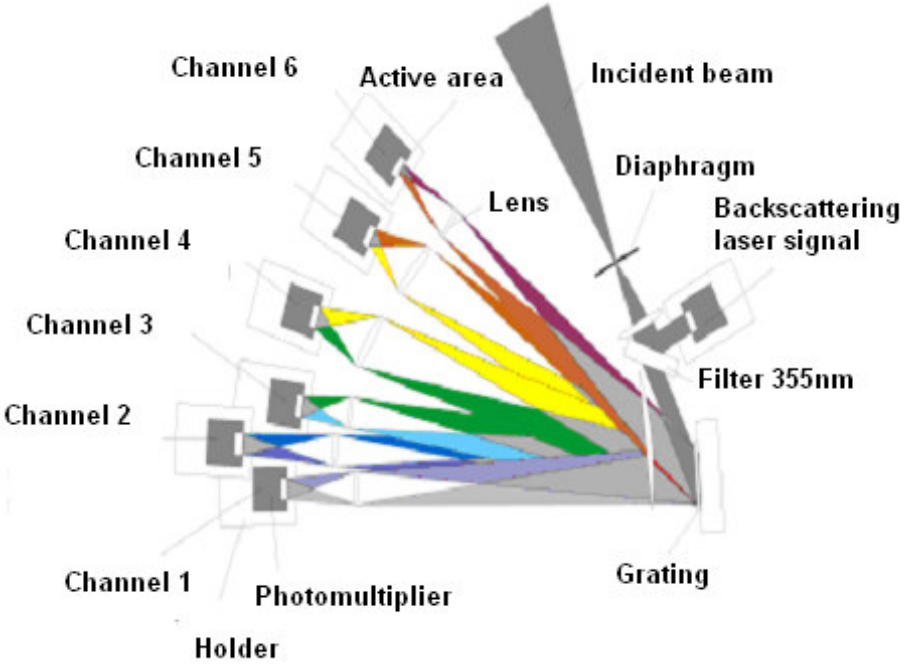


Figure 1: Optical design of the polychromator

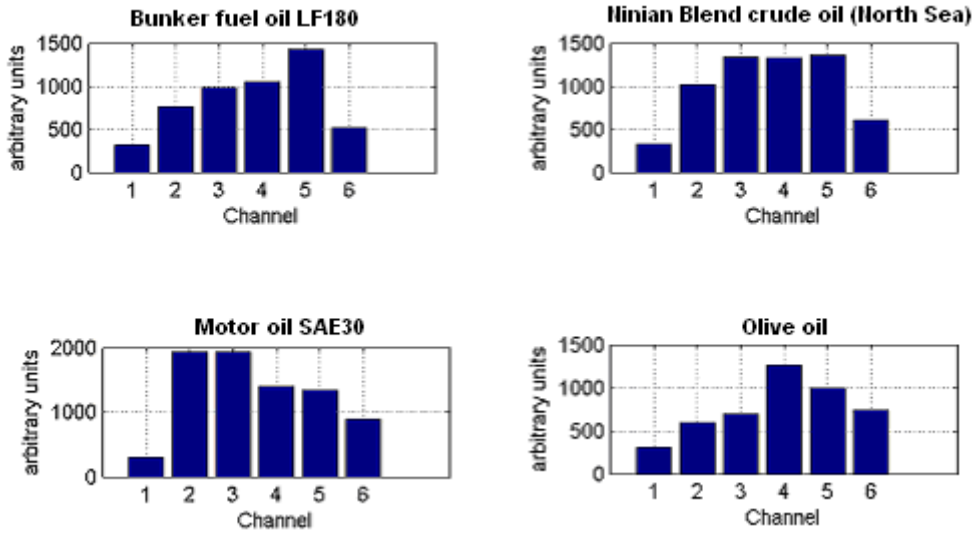


Figure 2: Examples of oil spectra spectrally grouped according to the detection channel wavebands of the laser fluorosensor.

Channel Relationship Method (CRM)

The channel relationship method (CRM) is simple method; it is a simple connection among the relative intensity of the normalized oil fluorescence detected by each channel of the polychromator. Some fluorescence signals have a strong appearance within a certain wavelength and a weak appearance within the other ranges of the wavelength. The fluorescence signals, which have shapes similar to each other, will form their own group in a diagram. We do expect that such crude oil where the fluorescence signatures are close to each other will distinguish themselves among the other oil classes. The same role will be applied to the rest of the oil classes with the exception of some oils from one or more classes that have close signatures to those from other oil classes,

which leads to difficulties in evaluating the class of such oils. CRM application was carried out using the same dataset as SVMs and MLP.

Artificial Neural Networks (ANN)

Artificial neural networks are powerful tools that can learn to solve problems in a way similar to how the human brain works. ANNs gather knowledge by detecting the patterns and relationships in data and learn (or: are trained) through experience, not from programming (14). 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 articial neural network where the output of such neuron is given by

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

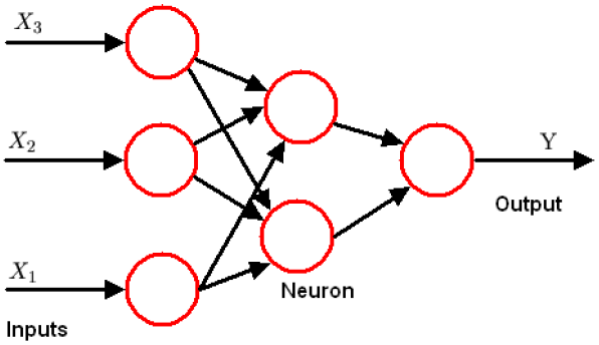


Figure 3: Structure of an Artificial Neural Network

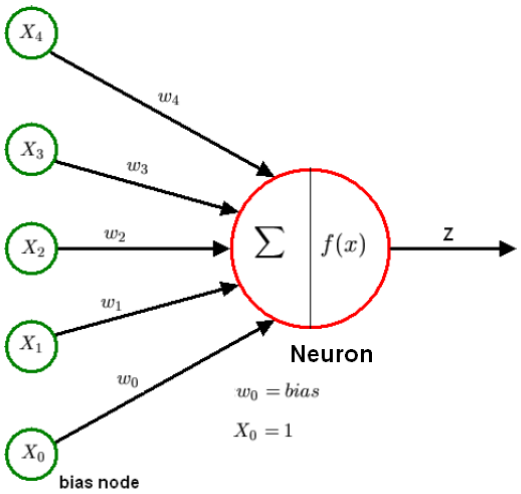


Figure 4: : Neuron Model of Artificial Neural Network

Many transfer functions could define $f(x)$ (Figure 5). The transfer fuction used in this work is Tanh function which is among the most poplar functions to the nueral network design due to its nice mathematical properties such as monotonicity, continuity, and differentiability , which are important to the training process [15].

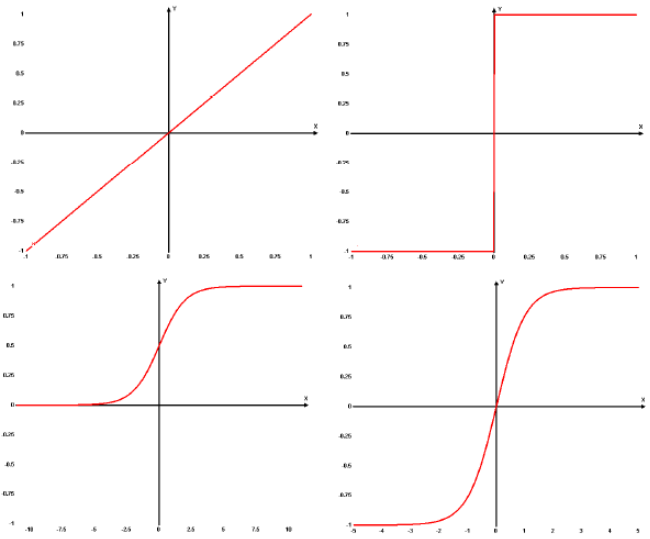


Figure 5: Transfer functions to neural network design

There are several types of ANNs according to their structure and learning algorithms. According to their structure ANNs can be classified as feedforward networks and recurrent networks (16). In a feedforward 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 (16). In recurrent networks, the output of some neurons is fed back to the same neurons or to neurons in a preceding layer (16). According to ANN learning process, the ANN can be classified to supervised learning, unsupervised learning, and 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 goodness of the neural network output corresponding to a given input (16).

In this paper the ANN shall estimate the unknown oil class in the laboratory dataset of the several 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 will be presented many times to the input of the forward activation place, and the process will continue until the actual outputs get closer to the desired output. The MLP that was used in this work contains three layers; input layer, one hidden layer and output layer (Figure 7). 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.

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

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

.

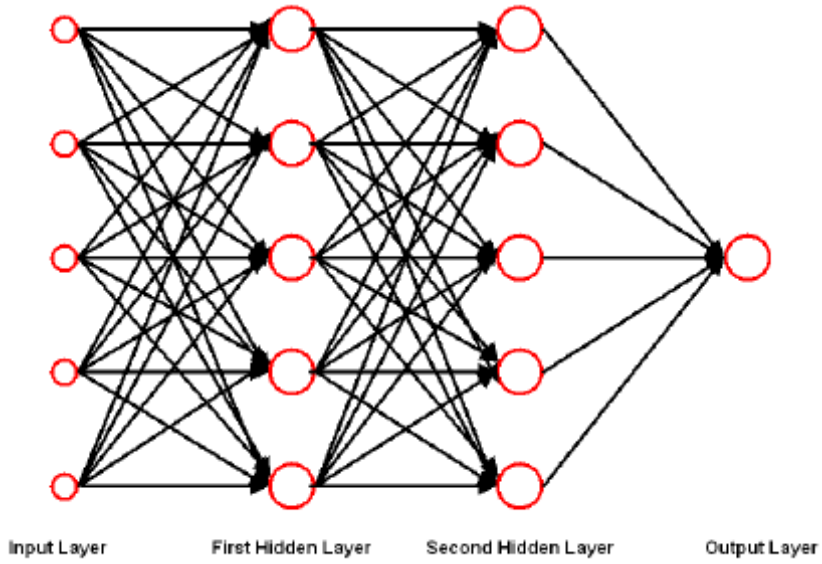


Figure 6: Multilayer Perceptron (MLP)

Figure 7 shows the MLP with the backpropagation algorithm that is used in this work. The components of this network are described in detail in reference (17). The network consists of an input layer (input Axon), non-linear hidden layer (hidden Axon), and output layer (output Axon). Hidden layer and the output layer apply tanh transfer function.

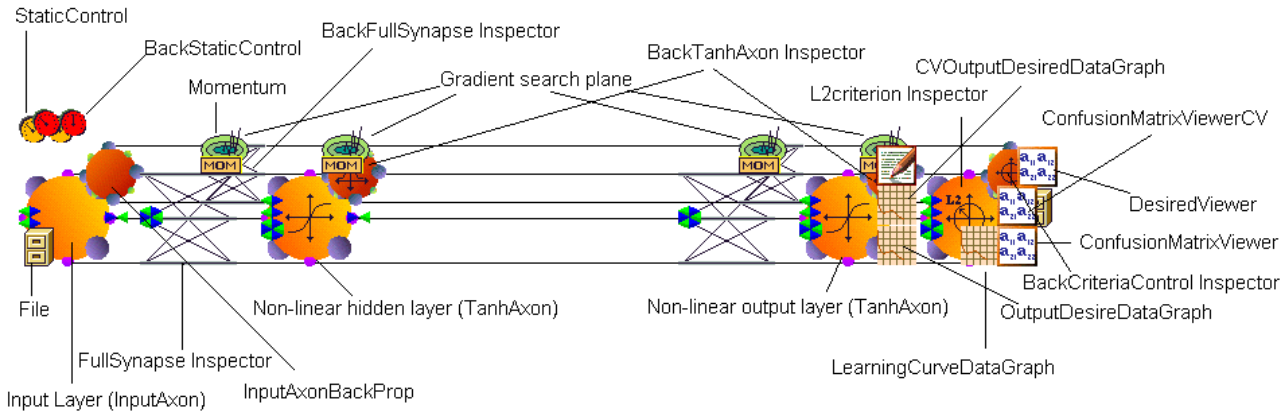


Figure 7: Multilayer Perceptron used to classify oil fluorescence. (More explanations needed. Is this original or taken from a website (then we must ask for a copyright permission)

The network has been trained based on 111 oil fluorescence signatures of crude oil, heavy oil, and sludge oil. The cross validation of the network uses a dataset of nine oil fluorescence spectra, with the aim to avoid over-training the network. The trained network then has been tested based on three oil fluorescence signatures. Then three unknown oil fluorescence spectra from the same three oil classes used in the training have been fed into the trained network for classification. The MLP was retrained in order to estimate the quality of the result and to investigate the stability of the trained MLP.

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 (18). Good explanation for SVMs theory and their applications is found in reference (19). The basic idea is to find the linear classifier called the hyperplane. Figure 8 shows many linear classifiers separate 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. This case is a simple case where the classes can be separated easily, this kind of support vector machines is called linear support vector machine (LSVM).

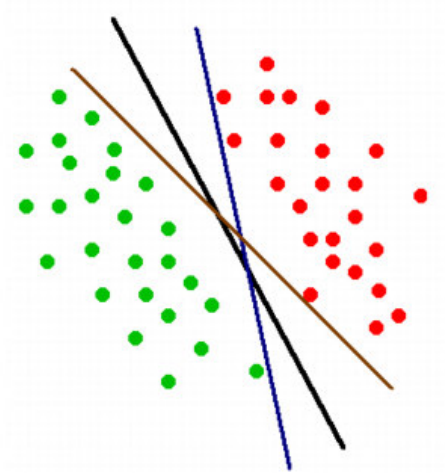


Figure 8: The best linear classifier (hyperplane).

For separable classes (Figure 9), a SVM classifier computes a decision function having a maximal margin M with respect to the two classes (blue and red classes. There are two planes touch the boundary of dataset, $wx+b=+1$ and $wx+b=-1$. w is vector perpendicular on the plane $wx+b=+1$. Let's consider x_2 to be any point on the plane $wx+b=+1$ and x_1 to be the closest point to x_2 on the plane $wx+b=-1$. The line from x_2 to x_1 is perpendicular to the planes. The maximum margin of the best classifier can expressed as $M = 2/\sqrt{w \cdot w}$.

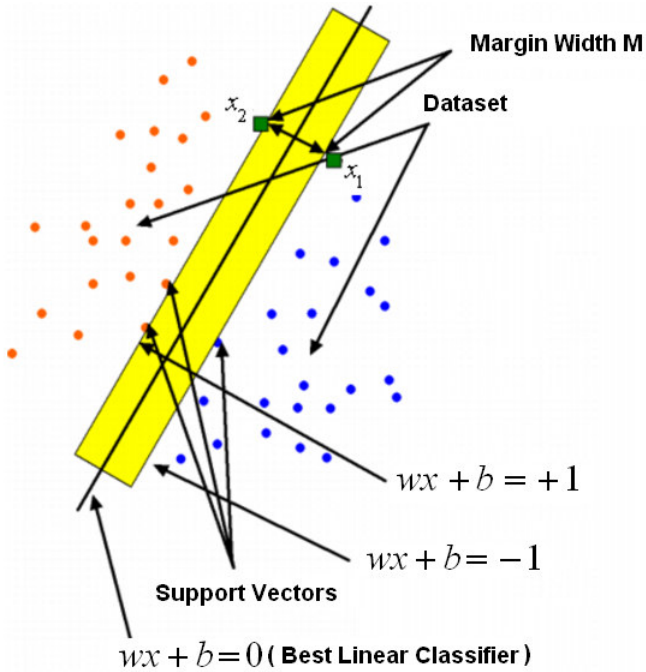


Figure 9: Classification of linearly separable dataset

Let's consider dataset: (x_1, y_1),....., (x_n, y_n), $\forall_i \in (1,.....,n)$, the decision boundaries can be found by solving the following constrained optimizing problem:

Minimize $\frac{1}{2}\|w\|^2$ subject to:

$$y_i(w \cdot x_i + b) - 1 \geq 0, \quad \forall_i \tag{4}$$

The Lagrange Function Formulation of this optimizing 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 Lagrange to zero:

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

This gives the conditions:

$$\sum_{i=1}^n \alpha_i y_i = 0 \quad and \quad w = \sum_{i=1}^n \alpha_i y_i x_i$$

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

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

subject to:

$$\alpha_i \geq 0, \qquad i = 1,.....,n \qquad \sum_{i=1}^n \alpha_i y_i = 0$$

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

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

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

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

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

In case of non linearly separable dataset (Figure 10), the a slack variable ζ_i is introduced leads to soft margin classifier:

Minimize $\frac{1}{2}\|w\|^2 + C\sum_{i=1}^n \zeta_i$ subject to:

$$y_i(w \cdot x_i + b) \geq 1 - \zeta_i \qquad \forall_i \tag{9}$$

The parameter C describes the trade-off between the maximal margin and the correct classification (19). These idea lead to the following dual problem

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

Subject to:

$$C \geq \alpha_i \geq 0, \qquad i = 1,.....,n \qquad \sum_{i=1}^n \alpha_i y_i = 0$$

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

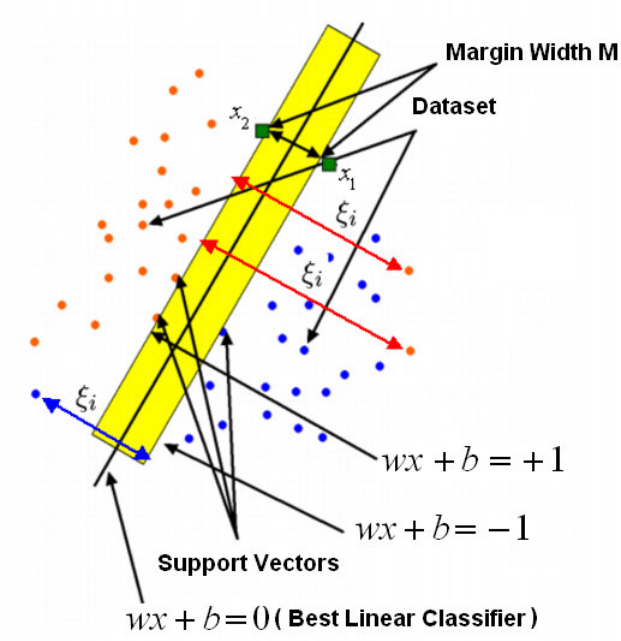


Figure 10:Classification of non linearly separable dataset

To solve non-linear classification problem, the linear support vector machines is applied to high dimensional space (Figure 11). 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. This means transforming from $\langle x_i, x_j \rangle$ to $\langle \Phi(x_i), \Phi(x_j) \rangle$ (Kernel trick $K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$

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

The common Kernel functions are Polynomial with degree d, Radial Base Function with width σ and Sigmoid with parameters k

$$K(x_i, x_j) = (\langle x_i, x_j \rangle + C)^d \quad (12)$$

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

$$K(x_i, x_j) = \tanh(k \langle x_i, x_j \rangle + \Theta) \quad (14)$$

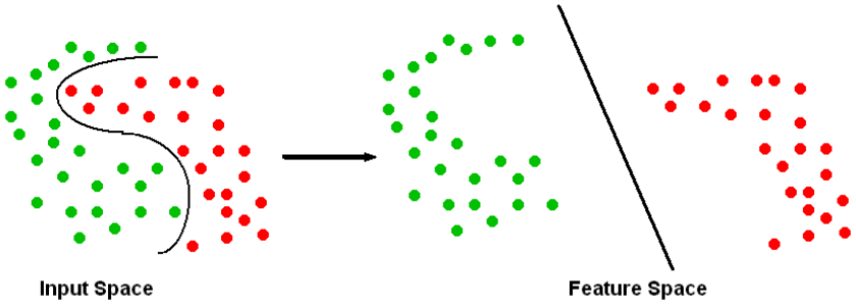


Figure 11: Transforming dataset from the input space to the high dimentional space

As in case of ANNs, NeuroSolutions was used in this work (17). In this program, transforming the data from input space to high dimensional space is done using a Radial Basis Function (RBF) network that places a Gaussian at each data sample (17). Thus, the feature space becomes as large as the number of samples (17). Figure 12 shows a SVM used to classify oil fluorescence spectral signature. The SVM was divided into two parts to implement the RBF dimensionality expansion and large margin classifier. As in the case of MLP, SVMs use the concept of backpropagation training to train the linear combination of guassians. SVMs are motivated by the concept of training and use only those inputs that are near the decision surface since they provide the most information about the classification (17). The components of the constructed SVM are explained in detail in reference (17).

Figure 12: Support Vector Machine used to classify oil fluorescence.

The training, testing, and production processes of SVM were carried out based on the same dataset that has been used in MLP to ensure an exact comparison in the quality of the results between the MLP and SVM.

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 has been achieved by two approaches. The first is to introduce noisy oil fluorescence data to be classified by the MLP and SVM that have been 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 with both approaches are compared.

Use of the noise-free dataset

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

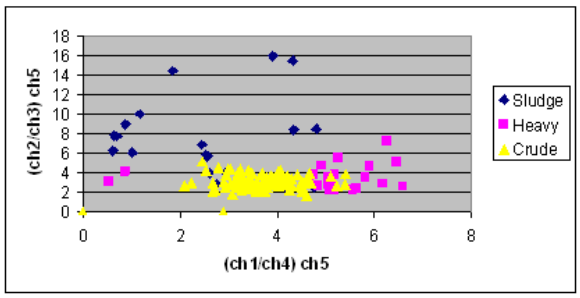
Training with noisy spectra

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 was added to them. I.e.,, the original dataset has been doubled to include also noisy spectra with a random Gaussian noise of 10%. Then, the same new noisy spectra were presented to the trained MLP and SVM.

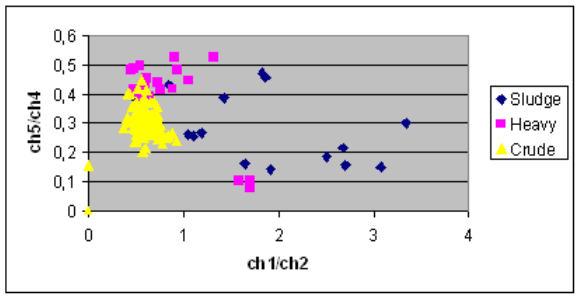
RESULTS AND DISCUSSION

Oil classification and identification using the CRM

Figure 13 shows the results of the oil classification using the CRM. Most of the oils within the crude class have shape fluorescence spectral signatures close to each other, 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 due to the oils within this class having a different shape oil fluorescence spectra than one another. It is shown also that some oils from the heavy and crude oils have close spectra to each other, which makes them mixed within the two classification groups.



(a)



(b)

Figure 13: Oil classification using the CRM

Oil classification and identification using MLP

Figure 14 shows the learning curve of the trained network where one can see that a mean squared error for both training (T) and cross validation (CV) approach zero. Table 1 shows the testing results of the trained network. The output of the tested ANN shows that the closest value to 1 is for the sludge oil for the first input and heavy oil is the second input and crude oil is the third input. This result meets the desired output performance and indicates the trained ANN meets the desired requirements. Table 2 shows the identification of unknown oil spectra by the trained network. They were three new fluorescence spectral signatures of oil, which the trained ANN 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 based on which class had the closest value to 1. Based on the database the results achieved by the trained ANN are 100%t correct.

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 based on the same dataset. This means that one should train the network many times to evaluate the best performance of the trained network.

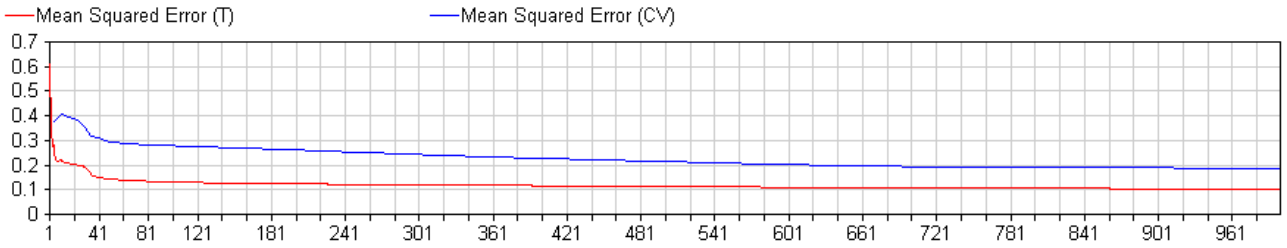


Figure 14: Learning curve of the trained MLP.

Table 1: Testing results of the trained network: desired and output. (more details on this would be good)

Des Sludge	Des Heavy	Des Crude	Out Sludge	Out Heavy	Out Crude
1.000000000000	0.000000000000	0.000000000000	0.862586719695	0.279102908007	0.001307395731
0.000000000000	1.000000000000	0.000000000000	0.022191138039	0.844636621374	0.228211478247
0.000000000000	0.000000000000	1.000000000000	0.052080805685	-0.017611388535	0.949783715848

Table 2: The identification of unknown oil spectra by the trained network.

Out Sludge	Out Heavy	Out Crude
0.869992051944	0.226982415077	-0.025177912220
0.083894094941	-0.006650917969	0.946461550285
0.036272837775	0.701901881172	0.329185320640

The following results are those of the re-trained network:

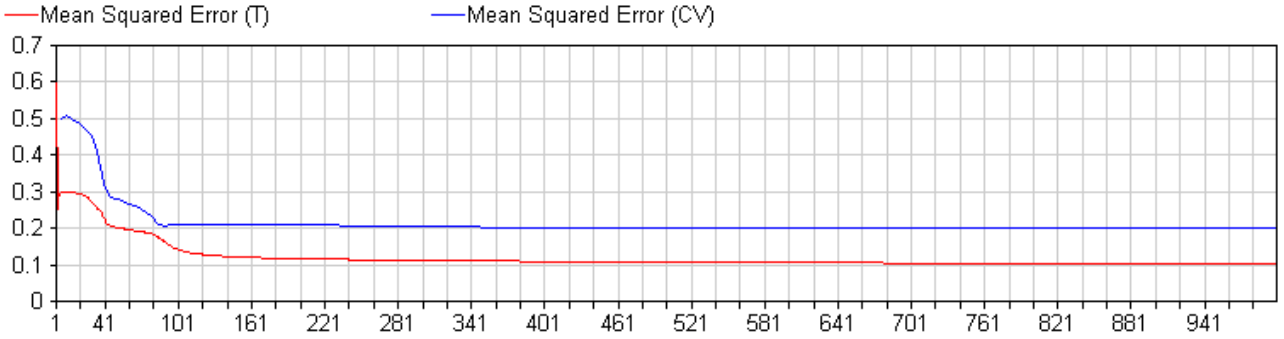


Figure 15: Learning curve the re-trained MLP.

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

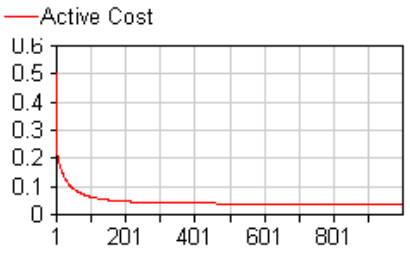
Des Sludge	Des Heavy	Des Crude	Out Sludge	Out Heavy	Out Crude
1.000000000000	0.000000000000	0.000000000000	0.782921004727	0.205500988145	-0.011546204816
0.000000000000	1.000000000000	0.000000000000	-0.012597677786	0.751779314508	0.232758694399
0.000000000000	0.000000000000	1.000000000000	-0.002296719375	-0.045376779999	0.992435021206

Table 4: The identification of unknown oil spectra by the re-trained network.

Out Sludge	Out Heavy	Out Crude
0.858248321504	0.270733718695	-0.030307910624
0.020858757939	-0.043168001660	0.978565838363
0.012214631381	0.634676111343	0.338985426887

Oil classification and identification using a SVM

Figure 16 (a) shows the active cost of SVM. From the comparison between the learning curves of MLP with the active cost curve of SVM, one can see that SVM 's curve approaches to zero more than those in case of MLP which means that SVM did understand the problem bettern than MLP. Tables 5 shows the testing results of the trained SVM. The output of the tested SVM shows the closest value to 1 is for the sludge oil for the first input and heavy oil is the second input and crude oil is the third input. This result meets the desired output performance and indicates the trained SVM meets the desired requirements. Table 6 shows the results of the identification by using SVMs. These results have been achieved from the first time training. The re-training process has no affect on the result. This means the performance of the SVM is achieved from the first time training process. This will save the user's time. The identification of an unknown oil fluorescence spectra is correct 100 percent. These identifications were made based on which class had the closest value to 1, the same role was applied in case of MLP. In terms of the accuracy with the comparison to MLP one can see that both MLP and SVMs have 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 16: (a) the active cost curve approaches zero, which means that classification of the dataset has been 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 SVMs.

Des Sludge	Des Heavy	Des Crude	Out Sludge	Out Heavy	Out Crude
1.000000000000	0.000000000000	0.000000000000	0.902840145617	0.065072020203	0.099350406861
0.000000000000	1.000000000000	0.000000000000	0.124211801782	0.817688061347	0.163424140172
0.000000000000	0.000000000000	1.000000000000	0.046857650318	0.060442064160	0.946875141582

Table 6: The identification of unknown oil spectra by a SVM

Out Sludge	Out Heavy	Out Crude
0.844488650966	0.103294605448	0.133288184330
0.052594011468	0.071544881890	0.937769266461
0.047724695721	1.097414896434	-0.054288792712

Classification of noisy oil spectra

Use of the noise-free dataset

Tables 7 to 10 show the identification of the 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 a crude oil and the third one as heavy oil. This identification is correct according to of the used types of oils.

Table 7: The identification of unknown noisy spectra (5% random Gaussian noise) by the trained MLP. The used spectra were from sludge, crude and heavy oil.

Out Sludge	Out Heavy	Out Crude
0.933746940830	0.182396993439	0.120517891761
0.028992114678	-0.047697817929	0.997649024759
-0.041814628213	0.794396639903	0.194699130363

Table 8 : The identification of unknown noisy spectra (10% random Gaussian noise) by the trained MLP. Same spectra as in Table 1.

Out Sludge	Out Heavy	Out Crude
0.891470276986	0.266454445878	0.069982940203
0.044947259027	-0.047771369003	0.995890799708
-0.036538975282	0.604728218534	0.337124831206

Table 9: The identification of unknown noisy spectra (5% random Gaussian noise) by the trained SVM. Same spectra as in Table 1.

Out Sludge	Out Heavy	Out Crude
0.681844599740	0.186857840115	0.266462216405
0.287588815174	0.276674914389	0.627709549873
0.282125561193	0.508499581877	0.438705740901

Table 10: The identification of unknown noisy spectra (10% random Gaussian noise) by the trained SVM. Same spectra as in Table 1.

Out Sludge	Out Heavy	Out Crude
0.651018139340	0.202725614300	0.291612532154
0.187760504317	0.184236255699	0.760981072141
0.239546491292	0.665344413959	0.302857117251

Training with noisy spectra

Tables 11 to 14 show the identification of the new noisy oil spectra that were 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 results' quality.

Table 11: The identification of unknown noisy spectra (5% random Gaussian noise) by the trained MLP.

Out Sludge	Out Heavy	Out Crude
0.954775736717	0.007657615361	-0.054846202693
-0.012897836339	-0.038194457891	0.877667977156
-0.027019623113	0.845652299360	0.470281100456

Table 12: The identification of unknown noisy spectra (10% random Gaussian noise) by the MLP.

Out Sludge	Out Heavy	Out Crude
0.953192302492	0.001498352980	-0.054780453367
0.000483113022	-0.043187622163	0.894579594259
-0.028441554014	0.724638494601	0.602675105316

Table 13: The identification of unknown noisy spectra (5% random Gaussian noise) by the trained SVM.

Out Sludge	Out Heavy	Out Crude
0.874742766358	0.127890650234	0.092008965337
0.160248039888	0.209935075175	0.774270741153
0.389574392508	0.479669466022	0.419895450568

Table 14: The identification of unknown noisy spectra (10% random Gaussian noise) by the trained SVM.

Out Sludge	Out Heavy	Out Crude
0.877556399054	0.125099231738	0.090387356557
0.057198064730	0.089671739067	0.924681649648
0.324682796707	0.477106689592	0.446996097880

CONCLUSIONS

The CRM provides a meaningful classification of the available spectra according to the rough oil type. From the results, one can reach the conclusion 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. Although the general result is still correct the accuracy will change. 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. This result 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 one needs 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 demonstrated the ability to classify noisy oil fluorescence of 5% and 10% Guassian noise. To ensure the best performance of the trained MLP and SVM the training dataset should include noise-free and noisy fluorescences.

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, one can train them offline and then apply them in real time application.. As a result, we reached the conclusion that the SVM classification method is the most convenient method to classify oil spectra..For the futures to work and 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 the 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 many available programs of ANNs and SVMs, so it is not imperative to restrict the programme used in this work. One might use another program and compare the result's quality.

ACKNOWLEDGEMENTS

The classification and identification of the oil fluorescence spectra have been carried out using the NeuroSolution program (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. **A**rtificial 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 Ming Tian, Song-Can Chen, Yi Zhuang & Jia Liu, 2004. Using statistical analysis and support vector machine classification to detect complicated attacks., 2004. In: Proceedings of 2004 International Conference on Machine Learning and Cybernetics, 5: 2747-2752 DOI 10.1109/ICMLC.2004.1378327

9 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

10 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, 2005

11 Catalogue of Optical Spectra of Oils, 2005 (University of Oldenburg, Institute of Physics, Section Marine Physics)

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

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

14 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

15 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. ISBN 0-8194-5987-9

16 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: pp. 1-3

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

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

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