PHYSOR 2004 -The Physics of Fuel Cycles and Advanced Nuclear Systems: Global Developments
Chicago, Illinois, April 25-29, 2004, on CD-ROM, American Nuclear Society, Lagrange Park, IL. (2004)

Support Vector Machine in Classification of Positron Lifetime Spectra

Senada Avdic
University of Tuzla, Department of Physics, Univerzitetska 4, 75000 Tuzla, Bosnia and Herzegovina

Positron annihilation spectroscopy is used in the study of radiation induced defects in nuclear materials in a non-intrusive way. The positron can be trapped by defects and the number of exponential components in the positron lifetime spectrum is related to the number of defect states. This work concerns classification of positron spectra with respect to the number of spectral components by support vector machine (SVM). The SVM has not yet been investigated for positron spectra analysis. The SVMs use an optimized generalization. The SVM classifier has been constructed by training with the simulated positron spectral data with two and three spectral components. Tuning the hyperparameters, such as the generalization parameter C, has been done by using the 10-fold cross-validation error. The experimental spectra available from polymer materials have been analysed by the constructed SVM nonlinear classifier. Experimental data are classified as the class of positron spectra with three spectral components with accuracy of 95.4 %. The SVM calculations show that certain degree of misclassification tolerance can produce a solution with good expected generalisation.

KEYWORDS: radiation defects, positron annihilation spectroscopy, number of spectral components, SVM classifier

1. Introduction

Positron annihilation spectroscopy is a diagnostic tool in the study of radiation induced defects in nuclear materials. Defect spectroscopy based on positron annihilation has been applied to study defect creation in steel used as construction material in nuclear reactors and also as the method for characterising radiation damage of the pressure vessel [1]. The positron can be trapped by defects such as vacancy and open-volume defects in materials. The annihilation of positrons in condensed matter yields the lifetime or annihilation rate of the positrons, which is sensitive to the local electron density in place the annihilation occurs. The measured positron lifetime spectrum provides information on the microscopic free-volume fraction and size distribution. The number of exponential decay terms in the positron lifetime spectrum is related to the number of defect states [2].

The main goal of the positron data analysis is the determination of the trapping rates and thus defect concentrations. There are some existing numerical codes, based on some kind of statistical minimisation principle, developed for positron lifetime analysis with different algorithms. The lifetime and the intensity values of each of the exponential components in the spectrum can be extracted by deconvolution, background subtraction and then a non-linear fitting process. A new method based on the use of artificial neural networks (ANNs) for unfolding mean lifetime and intensity of the spectral components of positron lifetime spectra was suggested and tested recently [3]. The capabilities of the method to analyze experimental data are demonstrated in the case of three component spectra obtained from free-volume measurements of polymer materials. In order to

achieve flexibility and to be able to unfold spectra with unknown number of spectral components, the standard ANN module has been combined with a special classification model based on self-organising map (SOM) and learning vector quantization (LVQ) algorithms [4]. However, the mentioned algorithms with visualisation technique cannot give clear information on classification of spectral data in input space since the visualisation technique is related to 2D map of neurons.

Good generalization is very important to minimize classification errors or to avoid over-fitting data. The goal of the support vector machine (SVM), a universal constructive learning method, is to optimize generalization [5]. The SVM method has been applied to the positron spectra classification in order to reduce classification errors. Application of the SVM shows advantages especially in a range of spectral data with overlapping samples from different classes, i.e. spectra with the different number of spectral components. Positron spectra with two and three components have been chosen as two classes of spectral data. Accuracy of the positron spectra classification by the SVM classifier has been investigated through analysis of the experimental positron spectra available from polymer materials.

2. Application of the SVM for classification

2.1 Basics of the SVM method

The support vector machines are based on the statistical learning theory. Some classical problems such as multi-local minima, curse of dimensionality and overfitting in neural networks, seldom occur in support vector machines. Training data are mapped into a high-dimensional feature space using a set of non-linear basis functions defined a priori so that the optimal discriminating boundary between classes is a hyperplane. The mapping is carried out by means of a 'kernel' function. Training involves solving a constrained quadratic problem for which numerous standard methods are available. The solution of the learning problem implicitly contains a subset of the training data, i.e. the support vectors that are particularly informative and significant members of the data set. Unlike neural networks where the number of model parameters that require estimation grow exponentially with the number of input features, the dimension of the SVM optimization problem is equal to the number of training samples [6].

2.2 Application of the SVM

This work concerns the applicability of the support vector machines for classification of the available experimental spectra from polymer materials depending on the number of spectral components. The simulated positron lifetime spectra with two and three components have been generated by the Posgen numerical code with the aim of constructing the SVM classifiers for the positron spectra. Although the SVM implements the Structural Risk Minimization (SRM) principle and can generalise well, a careful choice of the kernel function is necessary to produce a classification boundary that is topologically appropriate. In the case where a linear boundary is not appropriate the nonlinear SV machines are designed to map the input space into high-dimensional space, called the feature space. A set of nonlinear transformation functions is defined a priori that can map the input vector into high-dimensional feature space. The solution of this problem is a linear classifier in this feature space that creates a nonlinear separating

hypersurface in the original input space. The SVM constructs an optimal hyperplane in this higher dimensional space.

Using the nonlinear transformation functions $g_j(\mathbf{x})$, j=1,…,m to create the feature, decision function becomes [7]:

$$D(\mathbf{x}) = \sum_{j=1}^{m} w_j g_j(\mathbf{x}) \tag{1}$$

where the number of terms in the summation depends on the dimensionality of the feature space. The zero-order 'threshold' term w_0 is dropped since it can be represented by including a constant basis function (i.e. $g(\mathbf{x})=1$) in the feature space. This decision function in the dual form is given by:

$$D(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i y_i H(\mathbf{x}_i, \mathbf{x}) \tag{2}$$

The inner product kernel H is a representation of the basis functions $g_j(\mathbf{x})$, j=1,…,m. It is known a priori and used to form a set of approximating functions. It can be noticed that a kernel function is a function in input space. For a given set of basis function $g_j(x)$, the inner product kernel H is determined by the sum:

$$H(\mathbf{x}, \mathbf{x}') = \sum_{j=1}^{m} g_j(\mathbf{x}) \cdot g_j(\mathbf{x}') \tag{3}$$

where m may be infinite.

For classification of nonseparable data, the decision function is given by:

$$D(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i^* y_i H(\mathbf{x}_i, \mathbf{x}) \tag{4}$$

The parameters α_i^*, i=1,…,n, are the solution for the following quadratic optimisation problem:

Maximise the functional

$$Q(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j H(\mathbf{x}_i, \mathbf{x}_j) \tag{5}$$

subject to constraints

$$\sum_{i=1}^{n} y_i \alpha_i = 0, \quad \frac{C}{n} \geq \alpha_i \geq 0, \quad i = 1,...,n \tag{6}$$

given the training data $(\mathbf{x}_i, y_i)$, i=1,…,n, an inner product kernel H and regularisation parameter C.

'Spline' and 'polynomials' kernels have been chosen for testing. K-fold cross-validation is an efficient technique for tuning SVM hyperparameters. The regularization parameter C is a hyperparameter which determines the tradeoff between minimizing the training error and minimizing model complexity. For each of kernels we used 10-fold cross-validation over the training set which means that 1/10th of the training data was used as a validation set and the process was repeated with non-overlapping rotations. The value of parameter C was estimated using the validation set. All SVM calculations have been performed by the SVM numerical code [8]. The training data set consists of about 300 simulated spectral data points. Average cross-validation accuracy is 90.9 % with parameter C=0.001 for a linear spline kernel and 89.6 % with parameter C=1 for a polynomial kernel of degree 2. The test error rates are evaluated using parameters estimated from cross-validation. Average accuracy on the test set of 2000 simulated spectral samples by the spline kernel is 85.3 % for the two-component spectra and 98.4 % for the three-component spectra. The linear spline kernel with C=0.001 has been chosen as the optimal kernel in the range of linearly nonseparable spectral data taking into consideration topology of classification boundary, number of support vectors and misclassification error.

Decision boundary between the training spectral data with two (class 2) and three spectral components (class 3), obtained by nonlinear classifier with C=0.001 and 15 support vectors is shown in Fig.1. Support vectors are indicated by extra circles. The margin has no constant width due to the nonlinear projection into the input space.

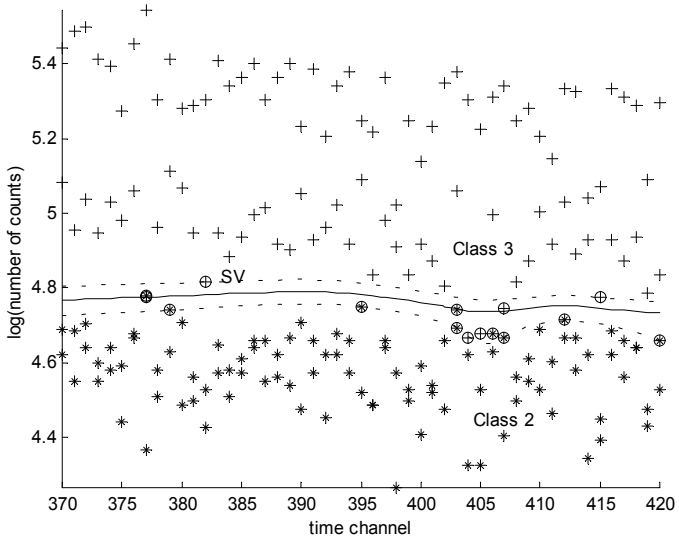

Fig.1 Decision boundary produced by the SVM using a linear spline kernel

The experimental spectra available from polymer materials have been analysed by the constructed nonlinear SV classifier. Experimental data containing of about 1600 points have been classified as the class of positron spectra with three spectral components with accuracy of 95.4 %.

3. Conclusion

The SVM calculations show that certain degree of misclassification tolerance can produce a solution with good expected generalisation. The applied SVM method provides more accurate classification information on the positron spectra compared to the ANN algorithms, especially for the overlapping classes. We demonstrated that the SVM with a linear spline kernel is an effective tool in discriminating the positron spectra regarding the number of spectral components.

Acknowledgements

The author wish to thank Dr. Marcus Schmidt for making available the experimental spectra for the ANN and SVM study.

References

1) J. Cizek, I. Prochazka, J. Kocik and E. Keilova, "Positron lifetime study of reactor pressure vessel steels", Phys. Stat. Sol. A178, 651 (2000).
2) A. Dupasquier and A. Mills, "Positron spectroscopy of solids", IOS Press, Amsterdam (1994).
3) I.Pazsit, R.Chakarova, P.Linden and F.H.J Maurer, "Unfolding positron lifetime spectra with neural networks", Applied Surface Science, 149, 97 (1998).
4) S.Avdic, R.Chakarova and I. Pazsit, "Analysis of the experimental positron lifetime spectra by neural networks", Nuclear Technology & Radiation Protection, Vol. XVIII, No.1, (2003).
5) V. Vapnik, "The nature of statistical learning theory", Springer, N.Y., (1995).
6) C.J.C. Burges, "A tutorial on support vector machines for pattern recognition", Data Mining and Knowledge Discovery, (1998).
7) V.Cherkassky, F.Mulier, "Learning from data", John Wiley & Sons, Inc., (1998).
S.Avdic, "Numerical code for the SVM classification of positron lifetime spectra", University of Tuzla, Department of Physics, unpublished paper, (2003).