Fibers and Polymers 2011, Vol.12, No.4, 546-549 DOI 10.1007/s12221-011-0546-x

Prediction of Rotor Spun Yarn Strength Using Support Vector Machines Method

Deogratias Nurwaha* and Xinhou Wang[1]

College of Textiles, Donghua University, Shanghai 201620, China
[1]*Key Laboratory of Science & Technology of Eco-Textile, Ministry of Education, China*
(Received November 20, 2010; Revised January 6, 2011; Accepted January 8, 2011)

Abstract: A new method for rotor spun yarn prediction from fiber properties based on the theory of support vector machines (SVM) was introduced. The SVM represents a new approach to supervised pattern classification and has been successfully applied to a wide range of pattern recognition problems. In this study, high volume instrument (HVI) and advanced fiber information system (Uster AFIS) fiber test results consisting of different fiber properties are used to predict the rotor spun yarn strength. The results obtained through this study indicated that the SVM method would become a powerful tool for predicting rotor spun yarn strength. The relative importance of each fiber property on the rotor spun yarn strength is also expected. The study shows also that the combination of SVM parameters and optimal search method chosen in the model development played an important role in better performance of the model. The predictive performances are estimated and compared to those provided by ANFIS model.

Keywords: SVMs, Yarn strength, Rotor spun yarn, Properties of fiber

Introduction

The relationship between fiber properties and yarn properties has been the focus of research in textile engineering. The development of the yarn strength prediction models from fiber properties have long been regarded as an important and widely studied issue in the textile engineering. The yarn strength prediction may have significant impact on controlling yarn quality. Early studies of yarn strength prediction used statistical techniques such as linear regression analysis have been done. Several studies using the neuro-fuzzy approach which takes the advantages of both neural networks and fuzzy logic to predict problems in textile are available [1-6]. Recently, a novel algorithm, called support vector machine (SVM), was developed by Vapnik and his co-workers [7]. Unlike most of the traditional neural network models which implement the empirical risk minimization (ERM) principle, SVM implement the structural risk minimization (SRM) principle which seeks to minimize an upper bound of the generalization error rather than minimize the training error. This induction principle is based on the fact that the generalization error is bounded by the sum of the training error and a confidence interval term that depends on the Vapnik-Chervonenkis (VC) dimension. There are no many examples of applications of SVM method available in textile issues. More recently, some studies using support vector machines in textile engineering have been reported [8,9]. As some previous studies have shown the nonlinear relationship between yarn strength and fiber properties [1-3], the SVM model can be a tool to overcome some difficulties in textile engineering. In this study, we will use the Support Vector Machines method as a new approach to predict the rotor spun yarn strength and the results will be compared to those from ANFIS model which has been cited as a powerful prediction model in more recent years [3-6].

*Corresponding author: jeadeo163@yahoo.com.cn

Introduction to the Support Vector Machines (SVM)

Support vector machines (SVM) are a group of supervised leaning methods that can be applied to classification or regression. Support vector machines represent an extension to nonlinear models of the generalized portrait algorithm developed by Vladmir Vapnik [7]. The SVM algorithm is based on the statistical learning theory and the Vapnik-Chervonenkis dimension introduced by Vladmir and Alexey Chervonenkis. A support vector machine performs classification by constructing an N-dimension hyper plane that optimally separates the data into two categories. Support vector machines models are closely related to neural networks. A Support Vector Machines (SVM) performs classification by constructing an N-dimensional hyperplane that optimally separates the data into two categories. Using a kernel function, SVM are an alternative training method for polynomial, radial basis function and multi-layer perceptron classifiers in which the weights of the network are found by solving a quadratic programming problem with linear constraints, rather than by solving a non-convex, unconstrained minimization problem as in standard neural network training.

The optimal plane classifier uses only dot products between vectors in input space. So the goal of SVM modeling is to find the optimal hyperplane that separates clusters of vector in such a way that cases with one category of the target variable are on one side of the plane and cases with the other category are on the other size of the plane. The vectors near the hyperplane are the support vectors. The Figure 1 presents an SVM process in the case of a two-dimensional example when the data has a categorical target variable with two

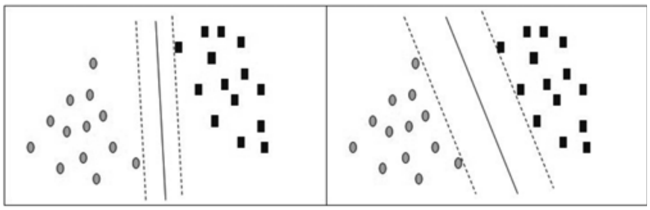

Figure 1. SVM of a two-dimensional example.

categories. One category of the target variable is represented by rectangles while the other category is represented by ovals.

The hyper plane can be constructed by solving a convex optimization problem that minimizing a quadratic function under linear inequality constraints. The optimization problem used to find the optimal hyper plane and the decision function used for the actual classification of vectors can be expressed in dual form from which depend only on dot products between vectors. The dual representation of the decision function is:

$$f(X) = \mathrm{sgn}\left(\sum_{i=1}^{l} Y_i \alpha_i < X \cdot X_i > + b\right) \tag{1}$$

Where $\alpha_i \in R$ is a real-valued variable that can be viewed as a measure of how much information al value X_i has. Thus for vectors that do not lie on the margin value will be zero. The optimal hyperplane classifier uses only dot products vectors in input space. In feature space this will be translate to $< \phi(X), \phi(X) >$. A kernel function $K(X, X')$ that gives two vectors in input, returns the dot product of their images in feature space is given by:

$$K(X, X') = < \phi(X), \phi(X') > \tag{2}$$

If we take equation (1) which is the decision function for the optimal hyper plane classifier in dual form and apply the mapping ϕ to each vector, we get

$$f(X) = \mathrm{sgn}\left(\sum_{i=1}^{l} Y_i \alpha_i < \phi(X) \cdot \phi(X_i) > + b\right) \tag{3}$$

The explicit mapping to feature space is not desirable. We will therefore use kernels which will give a non linear decision function of the form:

$$f(X) = \mathrm{sgn}\left(\sum_{i=1}^{l} Y_i \alpha_i K(X, X_i) + b\right) \tag{4}$$

The SVM algorithm is based on statistical learning theory, while being practical since it reduces to an optimization problem with a unique solution. A generalization to regression, that is, having $y \in R$, can be given. In this case the algorithm tries to construct a linear function in the feature space such that the training point lies to a distance of $\varepsilon > 0$. Similar to the pattern-recognition case, this can be written as a quadratic programming problem in terms of Kernels. The kernel approach is employed to address the curse of dimensionality. The support vector regression solution, using an ε-insensitive loss function is given by:

$$\max_{\alpha, a^*} W(\alpha, \alpha^*) = \max_{a, a^*} \sum_{i=1}^{l} \alpha_i^*(y_i - \varepsilon) - \alpha_i(y_i + \varepsilon)$$

$$-\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)K(X_i, X_j) \tag{5}$$

with constraints: $0 \le \alpha_i$, $\alpha_i^* \ge C$, $i = 1, \ldots, l$, α_i, α_i^* are the Lagrange multipliers

$$\sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0 \tag{6}$$

The regression equation is given by:

$$f(X) = \sum_{SVs}(\bar{\alpha}_i - \bar{\alpha}_i)K(X, X_i) + \bar{b} \tag{7}$$

Where

$$< \bar{w}, X > = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(X, X_i) \tag{8}$$

$$\bar{b} = \frac{1}{2}\sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(X_i, X_r) + K(X_i, X_s) \tag{9}$$

The equality constraint may be dropped if the Kernel contains a bias term, b being accommodated with the Kernel functions and the regression function is given by:

$$f(X) = \sum_{i=1}^{l}(\bar{\alpha}_i - \bar{\alpha}_i^*)K(X, X_i) \tag{10}$$

The quadratic loss function produces a solution which is equivalent to ridge regression, zeroth order regularization parameter $\lambda = 1/2C$ where C and γ are the optimization parameters. In this study, the following kernel function was used:

$$K(x, y) = e^{-\gamma|x-y|^2} \tag{11}$$

Data Collection

In order to get sufficient quantities of data included all fiber properties involved in this study, 1997-cotton crop study data from Texas Cotton Quality Evaluation studies at the international Textile Center were used in our study [10]. The fiber properties such fiber strength (S) in g/tex, micronaire (M) in μg/in., upper half mean length (UHML) in inches, fiber elongation (E), uniformity index (U) in % and yellowness (Y) in (+b yelloness) and grayness (G) in (Rd reflectance) measurements were made using the high volume instrument (HVI); short fiber content (SFC) in % measurements were made using Uster AFIS. All the spun yarns have been produced on an opened spinning machine with 30/1 count. Fiber properties were used as inputs (independent variables). Since multiple yarn sizes were spun from each cotton bale sampled, the yarn number (YC) was also included as an input variable. Count strength product (CSP) in lb×Ne was the target. We have divided the data into two sets, namely, training (estimating) and checking (validation) data sets.

Practical Considerations and Implementation

As the problem is nonlinear, we first apply the RBF kernel function to map the data into a different space where a hyper plane can be used to do the separation. The RBF function nonlinearly maps samples into a higher dimensional space, so it, unlike the linear kernel, can handle the case when the relation between class labels and attributes are nonlinear. Furthermore, the linear kernel and sigmoid kernel behave like RBF for certain parameters [11,12]. To find the optimal parameter we used grid and pattern search methods on C and γ. As the complexity increases by the number of support vectors, SVM is constructed through trading off decreasing the number of training error and increasing the risk of overfitting the data. Since SVM captures geometric characteristics of feature pace without deriving weights of network from the training data, it is capable of extracting the optimal solution with the small training set size. We conducted grid search and pattern search methods using four and ten-fold cross-validations on the training data and reported the validation results. One subset is chosen for testing and remaining 9 subsets are used for training and the process is repeated until all the subsets are chosen for the testing. For implementation, the DTREG software was used to execute the SVM.

Importance of Selection of the Optimization Methods and Number of Fold Cross-validation

We used both the grid search method on C=11634.5908 and γ=0.027144 and the pattern search method on C= 14718.1813 and γ=0.035152 using 4 and 10-folds cross-validation respectively. The results including the number of support vectors and analysis run time are reported in Table 1. We can look at the differences in results obtained with different types of optimal search methods, number of folds cross-validation and different parameters. We see that the grid search methods gives good results and uses less support vectors than the pattern search but more analysis runtime. The results are likely better for 10-folds cross-validation than for 4-folds cross validation for the pattern search method, better for 4-folds cross-validation for the grid search method.

Importance of Fiber Properties on Rotor Spun Yarn Strength

By examining which variables are used to split nodes near the top of the tree, the model determines the most important variables. A summary ranking overall variable importance is shown in Table 2. As we can see from the Table 2, the rotor spun yarn strength is influenced to a greater or less degree by the fiber properties.

Results Analysis

The prediction performance is evaluated using the following statistical metrics, namely, the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and proportion of variance explained by model that gives the percentage of correctly prediction residues (%). The smaller the values of RMSE and MAE, the closer are the predicted CSP values to the actual CSP values. The values of optimal parameters for both gird and pattern searches have been found (C=

Table 3. Comparison of prediction performances between SVM and ANFIS results [3] models for rotor yarn strength

Statistical parameter	SVM model	ANFIS model
RMSE	120.03	98.57
MAE	82.87	79.80
MAPE	4.07	3.77

Table 1. Comparison of the results from different methods

OM	NCVF	NSV	RMSE	MAE	MAPE	PVE	ART
Grid search	4	27	125.61	85.69	4.21	57.36	3s
	10	27	125.46	87.13	4.28	57.46	3s
Pattern search	4	34	131.72	89.20	4.38	53.12	1.5s
	10	34	127.04	88.06	4.31	56.39	1.5s

OM: optimization method, NCVF: number of cross validation folds, NSV: number of support vectors, RMSE: root mean square error, MAE: mean absolute error, MAPE: mean absolute percentage error, PVE (%): proportion of variance explained by the model, ART: analysis of run time (in seconds).

Table 2. Importance (%) of fiber properties on the rotor spun yarn strength

Fiber properties	YC	S	U	E	UHML	SFC	Y	G	M
%	100	34.57	12.62	12.11	6.15	4.46	3.69	2.45	2.24

38875.847 and γ=0.03119321) and the number of support vectors NSV=33. The results are the following: RMSE= 120.03, MAE=82.87, MAPE=4.07. Compared to the results from Table 1, we can see that the SVM prediction performance has been improved by choosing the appropriate optimal parameters C and γ. Table 3 shows the comparison of the results from the support vector machines (SVM) method using both grid and pattern search methods with optimal values of parameters to those from the adaptive neuro-fuzzy inference systems (ANFIS) method which have been more used as a good predictive model in more recent years [3-6]. The results from both models are similar with a little difference.

Conclusion

In this study, the results show that, like ANFIS, SVM can also provide an effective method for yarn strength prediction. Since we know that the relationship between fiber properties and yarn strength is nonlinear, SVM should be a potentially better data analytic method that needs to be explored more in depth to assess the practical impact of fiber properties on the rotor spun yarn strength. The results obtained have shown the importance of the optimal search methods, number of folds cross-validation and different parameters but especially the role of optimal parameters C and γ in improving the SVM performance. We have seen that both grid search and pattern searches methods can provide good results depending on the selection of the optimization parameters and the number of folds-cross validation. The results of this study need to be repeated and compared against others from similar analysis in order to assess the acceptable importance of fiber properties on rotor pun yarn strength.

References

1. L. Cheng and D. L. Adams, *Text. Res. J.*, **65**, 495 (1995).
2. S. Sette, L. Boullart, and P. Kiekens, *Text. Res. J.*, **65**, 200 (1995).
3. N. Deogratias and W. X. Hou, *Fiber. Polym.*, **11**, 97 (2010).
4. N. Deogratias and W. X. Hou, *Fiber. Polym.*, **19**, 782 (2008).
5. B. Karthikeyan and M. Sztandera, *International Journal of Clothing Science and Technology*, **22**, 187 (2010).
6. Z. A. Malik, M. H. Malik, T. Hussain, and A. Tanwari, *Ind. J. Fiber and Text. Res.*, **35**, 210 (2010).
7. C. Cortes and V. Vapnik, Support-Vector Network, Machine Learning, **20**, 273 (1995).
8. A. Ghosh and P. Chatterjee, *Fiber. Polym.,* **11**, 84 (2010).
9. P. H. Yap, X. Wang, L. Wang, and K. L. Ong, *Text. Res. J.*, **80**, 77 (2010).
10. Texas Cotton Quality Evaluation of Crop of 1997, International Textile Center, Lubbock, Texas.
11. N. Cristianini and J. Shawe-Taylor, “An introduction of Support Vector Machines and other Kernel-based Learning Methods”, Cambidge University Press, 2000.
12. N. Aronszajn, Theory of Reproducing Kernel, Trans. Amer. Math. Soc., 1950.
13. M. D. Ethridge, J. D. Towery, and J. F. Hembree, *Text. Res. Inst.*, **52**, 35 (1982).
14. J. S. R. Jang and N. Gulley, “The Fuzzy Logic Toolbox for Use with MATLAB”, The MathWorks Inc., 1995.
15. D. J. McCreight, R. W. Feil, J. H. Booterbaugh, and E. E. Backe, “Short Staple Yarn Manufacturing”, pp.311-375, 1997.