Performance of Neural Networks for Predicting Yarn Properties Using Principal Component Analysis

R. Chattopadhyay,[1] Anirban Guha,[1] Jayadeva [2]

[1]*Department of Textile Technology, IIT Delhi, New Delhi 110016, India*
[2]*Department of Electrical Engineering, IIT Delhi, New Delhi 110016, India*

Received 22 May 2002; accepted 30 June 2003

ABSTRACT: In recent years, neural networks have been used as a tool for modeling an industrial process. An improvement in their performance may be expected either by divining more efficient training algorithms or by intelligently manipulating the data set. The second method is examined. The problem chosen is one of predicting the properties of cotton yarn from the fiber properties. When the input data are known to correlate with each other, principal component analysis can be used to improve the performance of neural networks. © 2003 Wiley Periodicals, Inc. J Appl Polym Sci 91: 1746–1751, 2004

Key words: neural networks; yarn; principal component analysis

INTRODUCTION

Several research groups have shown neural networks as being quite successful in predicting yarn properties from the fiber properties or process parameters. Ramesh et al.[1] predicted the strength and elongation of air-jet spun yarns from the yarn count, percentage of polyester in a polyester–cotton blend, and front and back nozzle pressures. Cheng and Adams[2] predicted the CSP (see Appendix A) of ring yarns from the fiber properties measured in a high volume instrument (HVI). Pynckels et al.[3] tried to predict a wide range of properties of ring and rotor yarns from the fiber properties and machine parameters. Cabeco-Silva et al.[4] predicted the tenacity and elongation of carded cotton yarns using neural networks. In the case of cotton, a high correlation exists between various properties, which is well known. Finer cotton is usually longer, stronger, and contains less extraneous materials (trash). Similarly, in the case of synthetic fibers, the tenacity, modulus, and elongation at break are usually correlated to some extent. Therefore, we thought it worthwhile to investigate whether this knowledge could be used to develop neural networks that would be better at predicting the yarn properties from the fiber properties that are known to be correlated.

Correspondence to: Dr. R. Chattopadhyay (rchat@textile.iitd.ernet.in).

Journal of Applied Polymer Science, Vol. 91, 1746–1751 (2004)
© 2003 Wiley Periodicals, Inc.

EXPERIMENTAL

Study on ring spinning process

There are four major commercially viable technologies in the world for spinning staple fibers, of which ring spinning is the most popular. The data for the ring spun cotton yarn used in this study were obtained from a reputable industry source. They pertained to 20 distinct batches of fiber that were processed over a period of 4 months and the corresponding yarn properties. The fibers were tested on an HVI to collect information about the 2.5% span length, uniformity ratio, fiber fineness, bundle strength, and trash content. In a spinning mill, these fibers are processed by a series of machines that open, clean, separate, align, and twist the fibers to form yarn. In the present work, this entire series of machines has been simulated by a single neural network. A feedforward neural network was trained with five fiber properties and the yarn count as input and six yarn properties as the output. A single neural network (with six output units) was used for predicting the yarn properties as depicted in Figure 1.

The terms used in Figure 1 are explained in Appendix A. Out of the 20 data sets available, 14 were randomly chosen for training the network and the remaining 6 were used as the test set. The entire data set was scaled to lie between -1 and $+1$. The hyperbolic tangent (tanh) was used as the activation function for all the units, and a backpropagation algorithm was used for training. The network was stabilized by 10,000 cycles. The trained network was able to predict the training set with almost 100% accuracy. The average errors on the test set are presented in Table I. Note

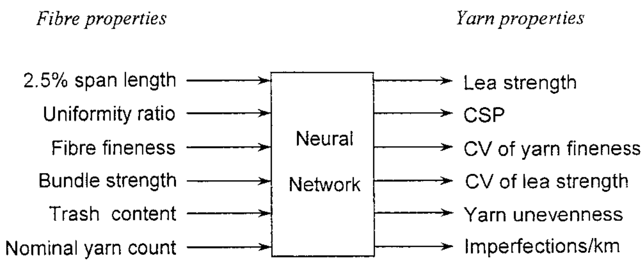

Figure 1 The structure of the network for predicting the properties of ring yarn.

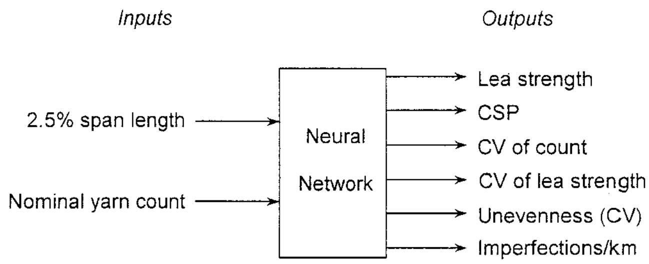

Figure 2 The structure of the truncated network for predicting the properties of the ring.

that the lea strength, count strength product, coefficient of variance (CV) of the strength, and yarn unevenness are very well predicted whereas the total imperfections per kilometer and CV of the yarn fineness are not. The average error is 7.5%.

One way of reducing the error is to reduce the complexity of the network by reducing the number of inputs. Because cotton properties are known to be correlated, an opportunity exists to reduce the number of fiber properties that are used as input. Therefore, we determined the correlation coefficients between the properties of the fibers used in the study and the results are shown in Table II.

The magnitudes of the correlation coefficients lie between 0.73 and 0.98 (see Table II). Except for two of them, all are greater than 0.8 in magnitude. Therefore, it was presumed that using only one of these five properties for the prediction of the six yarn properties may cause an improvement in the network's performance. This was expected because the information lost by neglecting the other four properties might be more than offset by the reduction in network size and the subsequent reduction in network complexity.

TABLE I
Average Error Percentages of Test Set for Predicting Properties of Ring Spun Cotton Yarn

Predicted property	Error (%)
Lea strength	3.9
Count strength product	2.7
CV of yarn fineness	19.1
CV of strength	3.4
Yarn unevenness (CV)	2.4
Total imperfections/km	13.6
Average	7.5

The data are from industry.

The choice of which fiber property to retain was made by looking at the correlation coefficients first and then with the knowledge about the HVI used for measuring the fiber properties. From Table II it can be seen that at least one correlation coefficient pertaining to the uniformity ratio, fiber fineness, and bundle strength is less than or equal to 0.8. Only the 2.5% span length and trash content have correlation coefficients above 0.8 with all the other properties in this case. Thus, our choice narrows down to a 2.5% span length and the trash content. It is also known that the HVI measures the trash content by optically scanning the surface of a fiber tuft and comparing the image with previously stored standard images from its database. This is an indirect method of measuring trash. The differences between these results and those obtained by the standard gravimetric method (i.e., by a Shirley Analyzer) are to be expected. On the other hand, an HVI measures the 2.5% span length by holding a fringe of fibers randomly and scanning it from one end to the other for the number of fibers. This comes quite close to the accepted method for measuring the span length. In the light of this discussion, the 2.5% span length was selected as the fiber property to be used to carry out the following exercise.

A feedforward neural network was trained with the same 14 data sets used earlier but using only the 2.5% span length and yarn count as inputs. The outputs were kept the same. The structure of the network is depicted in Figure 2.

TABLE II
Correlation Coefficients Among Properties of Fibers Used for Ring Spinning

	2.5% Span length	Uniformity ratio	Fiber fineness	Bundle strength	Trash content
2.5% Span length	1	0.87	−0.84	0.94	−0.93
Uniformity ratio	0.87	1	−0.73	0.80	−0.82
Fiber fineness	−0.84	−0.73	1	−0.91	0.93
Bundle strength	0.93	0.80	−0.91	1	−0.98
Trash content	−0.93	−0.82	0.93	−0.98	1

TABLE III
Comparison of Error Percentages of Test Set of Networks with 6 and 2 Inputs

Predicted property	Error (%) of networks with 6 Inputs	2 Inputs
Lea strength	3.9	3.4
Count strength product	2.7	4.0
CV of yarn fineness	19.1	20.0
CV of strength	3.4	7.8
Yarn unevenness	2.4	3.4
Total imperfections/km	13.6	16.6
Average	7.5	9.2

The training set error was almost zero. The errors of the test set (which was maintained as in the earlier exercise) are shown in Table III.

It can be seen from Table III that the prediction of the yarn properties, except the lea strength, deteriorated for the network with two inputs. In most of the cases, the deterioration was quite small except for the CV of the strength and total imperfections per kilometer. Nevertheless, the overall performance deteriorated from an error value of 7.5% to one of 9.2%.

These results show that reducing the number of inputs did not cause an improvement in the performance of a neural network, in spite of the fact that the inputs were highly correlated. This indicated that removal of the four fiber properties was resulting in a significant loss of information. It was therefore necessary to look for ways in which the inputs could be reduced (thereby reducing network complexity) without losing a significant amount of information. Principal component analysis provides a way of achieving this.

RESULTS AND DISCUSSION

Principal component analysis

The theory behind the development of principal component analysis can be found in many works.[5,6] Its goal is to transform a set of variables X_i ($i = 1, 2, \ldots, k$) into a new set of variables (P_i) called principal components, which are linear combinations of the Xs. These combinations are chosen so that the principal components satisfy two conditions.

1. the principal components are orthogonal to each other; and
2. the first principal component accounts for the highest proportion of total variation in the set of all Xs, the second principal component accounts for the second highest proportion, and so on.

Figure 3 shows a two-dimensional data set (plotted in the X_1–X_2 plane) that forms two clusters. The distribution of the data along the axes is also shown. Neither of these two axes can be termed more important than the other for describing the data. The principal components P_1 and P_2 can be drawn in such a way that the highest variation of the data occurs along P_1 (first principal component) and the next highest variance (in this case the lowest) occurs along P_2 (second principal component). It is now obvious that P_1 is more important for describing the data than P_2.

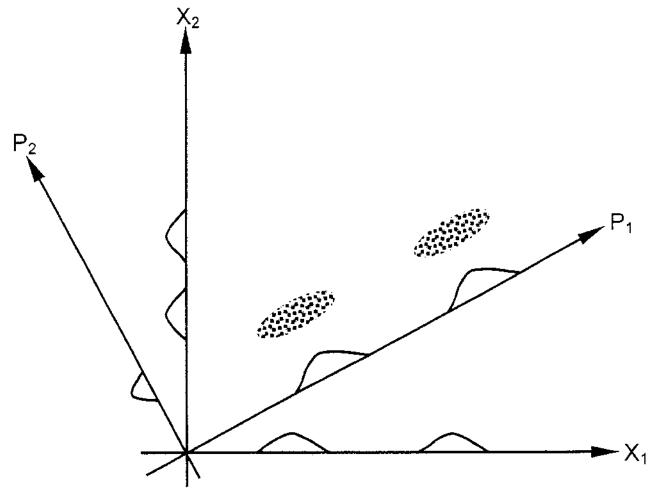

Figure 3 The principal components.

In the current problem, this technique needs to be applied to five-dimensional data. Once the relative importance of the five principal components is known, the least important components can be neglected. Now the projection of the data along the most important principal components will give a truncated data set with the least amount of information being lost. The steps involved in deriving the principal components from a data set are stated in Appendix B.

Application of principal component analysis

Using the procedure described in Appendix B, the principal components of the data set consisting of the five fiber properties were evaluated. The eigenvectors were the columns of the following matrix:

$$\begin{bmatrix} 0.8636 & 0.1023 & 0.2319 & -0.2908 & -0.3246 \\ -0.2733 & -0.1350 & -0.3843 & -0.8118 & -0.3169 \\ -0.2149 & 0.1788 & 0.7267 & -0.4656 & 0.4207 \\ -0.2058 & -0.6421 & 0.4962 & 0.1468 & -0.5269 \\ 0.3016 & -0.7260 & -0.1558 & -0.1345 & 0.5827 \end{bmatrix}$$

The corresponding eigenvalues were

$$(0.2602 \quad 0.1188 \quad 0.6961 \quad 1.5000 \quad 31.2348)$$

The projections of the original data on the eigenvectors resulted in the five fiber properties being trans-

TABLE IV
Comparison of Test Set Errors of Networks Trained with Original, Orthogonalized, and Orthogonalized Truncated Data

Predicted property	Error (%) of network with Original complete data	Orthogonalized complete data	Orthogonalized truncated data
Lea Strength	3.9	4.1	2.5
Count Strength Product	2.7	2.5	2.3
CV of yarn fineness	19.1	19.6	24.4
CV of Strength	3.4	3.1	7.8
Yarn unevenness	2.4	2.7	2.2
Total imperfections/km	13.6	13.0	3.5
Average	7.5	7.5	7.1

posed to another set of five properties that were orthogonal to each other. A neural network was created with the orthogonalized data as input. Next, by studying the eigenvalues, it was decided to retain only the first two principal components. The orthogonalized and truncated data (along with the yarn count) were also used to train a neural network for predicting the yarn properties. The errors of the test set compared with similar results for the network trained with the original values are shown in Table IV.

It can be observed that orthogonalization without any reduction of the input parameters did not cause any change in the overall performance of the network whereas orthogonalization with the reduction caused a slight improvement. Compared to the original six-input network, the three-input network gives lower prediction errors for four out of the six yarn properties. Of these, the improvements in the prediction of total imperfections per kilometer (13.6–3.5%) and lea strength (3.9–2.5%) are quite significant but those of the CSP and unevenness are marginal. The average error of all the properties improved from 7.5 to 7.1%.

This exercise indicated that the use of principal component analysis to orthogonalize and truncate a data set may allow the neural network's performance to improve. A reduction in the network size led to a reduction of the network complexity. Orthogonalization ensured the identification of the least important factors in the data and their subsequent removal. The end product was a smaller network with almost the same information content as the original network, and this led to an improvement of its performance.

Rotor spinning process

Rotor spinning technology is the second most popular staple fiber spinning technology in the world. Data pertaining to rotor yarns were obtained from industry, and they were used for conducting a similar exercise. In the original network, five fiber properties and the yarn count were used as inputs to a feedforward neural network to predict the six yarn properties (similar to the exercise carried out on the ring yarn data). Out of the 111 data sets available, 78 randomly chosen examples were taken as the training set and the remaining 33 were used as the test set. The errors of the trained network in the test set are shown in Table V.

It can be seen from the table that the lea strength, CSP, and unevenness (CV) are predicted very well whereas the CV of the strength, the CV of the count, and the total imperfections are very badly predicted. The average error is 17.4%.

Principal component analysis was carried out on the five fiber properties following the method described in Appendix B. The eigenvectors were the columns of the following matrix:

$$\begin{bmatrix} 0.0514 & 0.3264 & 0.3272 & -0.5913 & -0.6589 \\ 0.3443 & -0.8886 & 0.0131 & -0.2024 & -0.2253 \\ -0.8698 & -0.2769 & -0.0733 & 0.1474 & -0.3737 \\ 0.3309 & 0.1017 & 0.0634 & 0.7476 & -0.5632 \\ -0.1129 & -0.1297 & 0.9399 & 0.1697 & 0.2413 \end{bmatrix}$$

The corresponding eigenvalues were

$$(8.6777 \quad 12.9608 \quad 5.0479 \quad 2.4215 \quad 41.1866)$$

The projections of the original data on the eigenvectors resulted in the five fiber properties being transposed to another set of five properties that were orthogonal to each other. These orthogonalized data (along with the yarn count) were used to train a neural

TABLE V
Errors of Test Sets for Rotor Yarn Data

Yarn properties	Average error (%)
Lea strength	2.8
Count strength product	3.0
CV of yarn fineness	26.1
CV of strength	19.2
Yarn unevenness	3.1
Total imperfections/km	50.2
Average	17.4

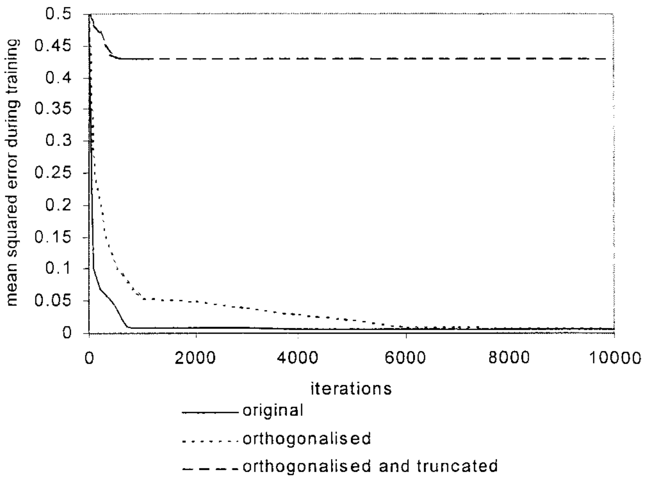

Figure 4 The training curves for three types of data for rotor yarn.

network for predicting the yarn properties. Next, a reduction of the least important orthogonalized components was attempted, as was done for the ring yarn data. Unfortunately, in this case, the network could not be trained at all. The training set errors were too high (25% average error) to conclude that proper training had been done. The training curves for the three types of rotor yarn data (original, orthogonalized, and orthogonalized and truncated) are shown in Figure 4.

The reason for the inability of the network to learn the orthogonalized and truncated data could be found in the correlation coefficients among the properties of the fiber used to spin the rotor yarns. These are shown in Table VI.

It can be seen that the correlation between the fiber properties is much weaker for the fibers used to spin rotor yarns than those used to spin ring yarns. In this case, except for one value, all the correlation coefficients are lower than or equal to 0.5 whereas for the ring yarn data, except for two, all were greater than 0.8. The low correlation coefficients could be attributable to the practice of adding an unspecified amount of soft waste from other mixings while spinning rotor yarns. Because of the low degree of correlation among the fiber properties, none of the orthogonalized components had an eigenvalue too low to be ignored. This is also reflected in the spread (ratio of largest to smallest) of the eigenvalues for ring and rotor yarn data. The spread of eigenvalues for the ring yarn data was 263 while that of the rotor yarn data was only 17. As a result, even the least important dimension in the orthogonalized rotor yarn data contained too much information to be ignored.

The errors in the test set of the orthogonalized (but not truncated) data compared with similar results for the network trained with the original values are shown in Table VII.It can be seen that the network with orthogonalized data as input gave lower prediction errors for all the yarn properties except CSP. In particular, for total imperfections per kilometer, the network with orthogonalized inputs gave a much lower error compared to the network with original data as input. The average error for all the yarn properties improved from 17.4 to 14.7%.

TABLE VII
Comparison of Error Percentages of Test Set for Predicting Rotor Yarn Properties

	Error (%) of networks with	
Predicted property	Original inputs	Orthogonalized inputs
Lea strength	2.8	2.7
Count strength product	3.0	3.2
CV of count	26.1	24.6
CV of strength	19.2	16.7
Unevenness (CV)	3.1	2.8
Total imperfections/km	50.2	38.0
Average	17.4	14.7

CONCLUSION

This study explored the possibility of improving the performance of neural networks. The following presumption was found to be incorrect: cutting down the number of inputs to a network, based on the strength of the correlation coefficients, would lead to better network performance because of a reduction in network complexity.

TABLE VI
Correlation Coefficients Among Properties of Fibers Used to Spin Rotor Yarn

	2.5% Span length	Uniformity ratio	Fiber fineness	Bundle strength	Trash content
2.5% Span length	1	0.17	0.50	0.83	−0.48
Uniformity ratio	0.17	1	0.30	0.32	−0.11
Fiber fineness	0.50	0.30	1	0.42	−0.27
Bundle strength	0.83	0.32	0.42	1	−0.52
Trash content	−0.48	−0.11	−0.27	−0.52	1

The improvement in the performance of neural networks was made possible by orthogonalizing the input data with the help of principal component analysis.

When the correlation between inputs was high, a reduction of the least important orthogonalized components could bring about a further improvement in the network's performance.

APPENDIX A

2.5% span length: the length above which 2.5% of the fibers lie when caught in a random manner

50% span length: the length above which 50% of the fibers lie when caught in a random manner

Bundle strength: the strength of a bundle of fibers held by two jaws at a distance of 1/8 in.

CSP: the product of the lea strength (lb) and yarn fineness in the English Cotton Count (number of strands in 120 yards of yarn that weigh 1 lb), which is useful for comparing the strengths of yarns of different finenesses

Imperfections/kilometer: the number of thin places, thick places, and neps present in every kilometer of yarn

Lea strength: a skein of 120 yards (consisting of 80 loops of yarn) is stretched to the breaking point and the force required is noted

Nep: an abnormal thick region in the yarn that contains 200% more mass than the average yarn for ring yarns (280% for rotor yarn) and has a length of less than 1 mm

Thin place: a region in the yarn that contains 50% less mass than the average yarn mass

Thick place: a region in the yarn that contains 50% more mass than the average yarn mass

Uniformity ratio: the ratio of the 50% span length and 2.5% span length (%)

Yarn unevenness: the CV of the weight of 8-mm pieces of yarn (obtained by passing 1 km of yarn between two parallel capacitor plates and noting the continuous change in capacitance)

APPENDIX B

Given a set of data, the principal components are the eigenvectors of the covariance matrix sorted in the decreasing order of the corresponding eigenvalues (i.e., the first principal component is the eigenvector corresponding to the largest eigenvalue). Let the data be arranged in the form of a matrix with m rows and n columns, where the rows indicate the samples and the columns indicate the properties. The following steps need to be performed to extract the principal components from the data.

Step 1: The data are first converted to a set of values with zero mean by subtracting the average of each column from each of the values of the column. Let any row of this zero-mean matrix be given by

$$y_1 \quad y_2 \quad \cdots \quad y_n$$

Step 2: The correlation matrix corresponding to this row must be calculated as follows:

$$\begin{bmatrix} (y_1)^2 & (y_1\, y_2) & \cdots & (y_1 y_n) \\ (y_2\, y_1) & (y_2)^2 & \cdots & (y_2\, y_n) \\ \vdots & \vdots & & \vdots \\ (y_n y_1) & (y_n\, y_2 & \cdots & (\mathrm{y}_n)^2 \end{bmatrix}$$

Step 3: The correlation matrices for all the m rows must be evaluated.

Step 4: All the correlation matrices obtained in step 3 must be added to attain the final correlation matrix.

Step 5: The eigenvalues and eigenvectors of this correlation matrix should be calculated. The eigenvectors are the principal components and the eigenvalues give their relative importance.

Step 6: The projection of the original matrix onto the principal components gives an orthogonalized data set of n dimensions. The relative importance of each dimension is given by the corresponding eigenvalue.

References

1. Ramesh, M. C.; Rajamanickam, R.; Jayaraman, S. J Text Inst 1995, 86, 459.
2. Cheng, L.; Adams, D. L. Text Res J 1995, 65, 495.
3. Pynckels, F.; Kiekens, P.; Sette, S.; Van Langenhove, L.; Impe, K. J Text Inst 1997, 88, 440.
4. Cabeco Silva, M. E.; Cabeco Silva, A. A.; Samarao, J. L.; Nasrallah, B. N. In Proceedings of the Beltwide Cotton Conference, Nashville, TN, 1996; Vol. 2, p 1481.
5. Haykin, S. Neural Networks: A Comprehensive Foundation; Macmillan: Englewood Cliffs, NJ, 1994; p 363.
6. Hertz, J.; Krogh, A.; Palmer, R. G. Introduction to the Theory of Neural Computation; Addison–Wesley: Redwood City, CA, 1991; p 204.