ESANN 2016 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 27-29 April 2016, i6doc.com publ., ISBN 978-287587027-8. Available from http://www.i6doc.com/en/.

Using Robust Extreme Learning Machines to Predict Cotton Yarn Strength and Hairiness

Diego P. P. Mesquita[1], Antonio N. Araújo Neto[1],
José F. Queiroz Neto[1], João P. P. Gomes[1] and Leonardo R. Rodrigues[2] *

1- Federal University of Ceará - Computer Science Department
Rua Campus do Pici sn - Fortaleza-CE - Brazil

2- Institute of Aeronautics and Space - Electronics Division
Praça Marechal Eduardo Gomes, 50, São José dos Campos-SP - Brazil

Abstract. Cotton yarn is often spun from a mixture of distinct cotton bales. Although many studies have presented efforts to predict hairiness and strength from cotton properties, the heterogeneity of this mixture and its influence in such values have been neglected so far. In this work the properties of the cotton bale mixture are modeled as random variables and a robust variant of the Extreme Learning Machine (ELM) to address the cotton quality prediction problem is proposed. A real world dataset collected from a textile industry was used to compare the performance of the proposed model with a traditional ELM and a linear regression model. The results showed that the proposed method outperformed the benchmark methods in terms of Average Root Mean Square Error (ARMSE).

1 Introduction

Predicting cotton yarn quality based on the analysis of the cotton fiber has been the objective of many studies in last years [1]. Being able to predict yarn quality may reduce production costs by allowing better production planning in textile industries. In recent years, several works have proposed the use of machine learning methods for this task [2, 3]. Artificial Neural Networks (ANN) have been used to predict different cotton yarn quality metrics [4]. Among these metrics, strength and hairiness are two of the most important ones.

Strength is an important mechanical property of cotton yarns which is related to the fiber structure [5]. The measurement of strength has become an important research topic over the years. Many factors may affect strength measurements such as environmental conditions (temperature and moisture), gauge length, processing history of cotton samples and timing of the test [6]. Hairiness is characterized by the quantity of freely moving fiber ends or fiber loops projecting from a yarn [7]. Usually hairiness is undesirable if it is too high and some common problems are breakages, lower machine efficiency in speed knitting and a bad appearance of the produced fabrics [8].

The cotton quality prediction problem can be defined as a standard regression problem where the inputs are features extracted from cotton fiber bales and the outputs are the cotton yarn quality metrics. Although good results have

*This work was supported by the Brazilian National Council for Scientific and Technological Development (CNPq), grant 402000/2013-7.

been obtained so far, previous works do not address many real world problems faced by many textile companies. One problem relies on the fact that in practice it is very uncommon to produce cotton yarns from a homogeneous (with similar properties) cotton mixture. Cotton yarns are produced instead from a mixture of different cotton bales, each with its own properties. As a consequence, the mapping between fiber properties and yarn quality can not be done in a straightforward way since there is not a single input feature vector, but different feature vectors obtained from the different bales. It is worth pointing that the variability of the input bales may influence the final cotton yarn quality.

This work proposes a Robust Extreme Learning Machine (R-ELM) to predict cotton yarn strength and hairiness under uncertain inputs. The input feature vectors are modeled as random variables that can be defined using the feature vectors of the bales that belong to the mixture. The algorithm is said to be robust since the output is robust to this uncertainty on the training set.

2 Proposed Method

2.1 Overview of Extreme Learning Machines

The ELM model was originally proposed in [9] as an alternative learning method for single hidden layer feedforward neural networks (SLFNs) using backpropagation. The main difference between a conventional SLFN and an ELM is that the latter assigns random values to the weights between the input layer and the hidden layer. Then, the weights between the hidden layer and the output layer can be computed using the Ordinary Least Squares (OLS) estimate.

Consider a training set containing n arbitrary distinct samples (x_i, y_i), where $x_i \in \mathbb{R}^{d\times 1}$ and $y_i \in \mathbb{R}^{m\times 1}$. Let $w_j \in \mathbb{R}^{d\times 1}$ be the weight vector connecting the j-th hidden node and the input nodes, b_j be the bias of the j-th hidden node and $g(\cdot)$ be the activation function. The output o_i of a SLFN with h hidden neurons for the i-th sample is then given by

$$o_i = \sum_{j=1}^{h} \beta_j g(w_j \cdot x_i + b_j) \tag{1}$$

where $\beta_j \in \mathbb{R}^{m\times 1}$ is the weight vector connecting the j-th hidden neuron and the output neurons. Equation 1 can be written in a compact form using Eq. 2:

$$H\beta = O \tag{2}$$

where $O = [o_1, o_2, \ldots, o_N]^T$ and the elements of matrix H are computed as defined in Eq. 3.

$$h_{ij} = g(w_j \cdot x_i + b_j) \tag{3}$$

The output weight vector $\beta = [\beta_1, \beta_2, \ldots, \beta_h]^T$ must be estimated in order to minimize the error function presented in Eq. 4. The OLS solution, given by Eq. 5, is used to find an estimate for β which minimizes Eq. 4.

$$min \quad ||H\beta - Y||^2 \tag{4}$$
$$\hat{\beta} = (H^T H)^{-1} H^T Y \tag{5}$$

The ELM approach can be summarized in the following steps:

- **Step 1:** Randomly assign values for each input weight vector w_j and each threshold b_j;
- **Step 2:** Compute the hidden layer output matrix H using Eq. 3;
- **Step 3:** Estimate the output weight β using the Moore-Penrose pseudoinverse according to Eq. 5:

2.2 Robust Extreme Learning Machines (R-ELM)

Consider a regression problem where the input data x_i are random variables. In the proposed approach, the goal is to obtain the ELM output given this uncertain input. For that, it is necessary to estimate the corresponding hidden layer output vector H_i. It is important to notice that, since x_i is a random variable, H_i can also be modeled as a random variable.

According to Eq. 3, H_i can be calculated by computing the nonlinear function $g(\cdot)$ over the inputs x_i, the hidden layer weights and the biases. It can be noticed that, since x_i is a random variable with unknown distribution, estimating the random variable H_i can be a non tractable problem. In such situation, a possible solution can be obtained by using Monte Carlo techniques.

In the current problem, the random inputs x_i are the feature vectors of each mixture of bales. Consider that the i-th mixture is composed of P bales, the feature vectors for each bale can be represented by $x_i^{(p)}$ where $p = (1, \ldots, P)$. In this work, x_i is modeled as $\mathcal{N}(\overline{x}_i, \Sigma^{(i)})$, where $\overline{x}_i$ and $\Sigma^{(i)}$ are calculated according to Eqs. 6 and 7.

$$\overline{x}_i = \frac{1}{P}\sum_{p=1}^{P} x_i^{(p)} \tag{6}$$

$$\Sigma^{(i)} = \frac{1}{P}\sum_{p=1}^{P} (x_i^{(p)} - \overline{x}_i)(x_i^{(p)} - \overline{x}_i)^T \tag{7}$$

The Monte Carlo procedure consists in sampling K points from the distribution of x_i. After that the hidden layer outputs, $H_i^{(k)}$ for $k = (1, \ldots, K)$ are calculated according to Eq. 3. After obtaining the K samples for a given H_i, any distribution estimation procedure (parametric or non-parametric) can be used.

Considering H as a random matrix with the n vectors H_i, it is possible to represent matrix H according to:

$$H = \overline{H} + U \tag{8}$$

where $\overline{H}$ is the mean matrix of H and U is a random matrix with zero mean.

The objective function presented in Eq. 4 can be redefined so that H is a random matrix. In this case, the goal becomes to minimize the expected value of the objective function as defined in Eq. 9.

$$min \quad E\left[||H\beta - Y||^2\right] \tag{9}$$

The objective function can be expressed as:

$$\begin{aligned}
\mathrm{E}[||H\beta - Y||^2] &= \mathrm{E}[((\overline{H} + U)\beta - Y)^T((\overline{H} + U)\beta - Y)] \\
&= \mathrm{E}[((\overline{H}\beta - Y)^T + (U\beta)^T)((\overline{H}\beta - Y) + U\beta)] \\
&= (\overline{H}\beta - Y)^T(\overline{H}\beta - Y) + E(\beta^T U^T U\beta) \\
&= ||\overline{H}\beta - Y||^2 + \beta^T \Sigma^H \beta
\end{aligned}$$

where $\Sigma^H = E[U^T U]$ is the covariance matrix of H.

As observed, this problem, termed as robust approximation problem, can be seen as a regularized least square problem and its solution is given by:

$$\hat{\beta} = (\overline{H}^T\overline{H} + \Sigma^H)^{-1}\overline{H}^T Y \tag{10}$$

It is possible to obtain the matrix $\overline{H}$ and the covariance matrix $C^{(i)}$ for the i-th transformed example by the hidden layer using Eqs. 11 and 12.

$$\overline{H}_i = \frac{1}{K}\sum_{k=1}^{K} H_i^{(k)} \tag{11}$$

$$C^{(i)} = \frac{1}{K}\sum_{k=1}^{K}(H_i^{(k)} - \overline{H}_i)(H_i^{(k)} - \overline{H}_i)^T \tag{12}$$

The covariance matrix Σ^H can be obtained by summing $C^{(i)}$ over all i examples, as shown in Eq. 13:

$$\Sigma^H = \sum_{i=1}^{n} C^{(i)} \tag{13}$$

This result can be obtained by noticing that for each element j, j' of Σ^H it follows that:

$$\begin{aligned}
\Sigma^H_{j,j'} &= \sum_{i=1}^{n}\mathrm{E}[U_{i,j}U_{i,j'}] = \sum_{i=1}^{n}\mathrm{Cov}(U_{i,j}U_{i,j'}) + \mathrm{E}[U_{i,j}]\mathrm{E}[U_{i,j'}] \\
&= \sum_{i=1}^{n}\mathrm{Cov}(U_{i,j}U_{i,j'})
\end{aligned}$$

The final procedure to implement the R-ELM model can be summarized in the following steps:

- **Step 1:** Calculate the mean vectors ($\overline{x}_i$) and the covariance matrices ($\Sigma^{(i)}$) using the feature vectors that belong to the i-th mixture according to Eqs. 6 and 7.
- **Step 2:** Generate K samples from each random variable x_i with distribution defined as $\mathcal{N}(\overline{x}_i, \Sigma^{(i)})$.
- **Step 3:** For the i-th example, obtain $H_i^{(k)}$ for all K samples, according to Eq. 3.
- **Step 4:** Calculate $\overline{H}$ and Σ^H using Eqs. 11, 12 and 13.
- **Step 5:** Find the weights β of the output layer according to Eq. 10.

3 Experiments

The proposed method was tested on a dataset of cotton yarn quality measures collected in 2009 on a facility of a textile industry in Brazil. A total of 640 mixtures of 10 bales were collected along with the corresponding strength and hairiness of the produced yarn. The feature vector for each bale comprises the following measurements: Trash Code, Trash Area, Trash Particle Count, Length, Uniformity, Short Fiber Index, Strength, Elongation, Micronaire, Maturity, White Level and Yellow Level.

The performance of the proposed algorithm was compared against a traditional ELM and a linear regression model. The experiments consisted of 20 similar trials with 90% of the dataset in the training set and 10% in the testing set. The number of hidden nodes for R-ELM and ELM was selected using 10-fold cross-validation. In order to assess the influence of the number of samples used to approximate each H_i, the R-ELM algorithm is evaluated for different values of K. The performance measure adopted was the Average Root Mean Square Error (ARMSE). The results obtained in the experiments are shown in Table 1.

Table 1: ARMSE comparison for cotton yarn hairiness and strength prediction

Model	**Hairiness**	**Strength**
Linear Regression	1.279 ± 0.157	2.732 ± 0.805
ELM	1.075 ± 0.217	2.484 ± 0.764
R-ELM ($K = 5$)	1.045 ± 0.227	2.355 ± 0.820
R-ELM ($K = 10$)	1.021 ± 0.222	2.333 ± 0.774
R-ELM ($K = 15$)	1.015 ± 0.223	2.307 ± 0.714
R-ELM ($K = 20$)	1.002 ± 0.223	2.297 ± 0.731
R-ELM ($K = 25$)	0.998 ± 0.225	2.228 ± 0.739

As expected, the proposed R-ELM model presented more reliable results than the other methods. The linear model, which is commonly adopted for this task, presented the worst performance. It is worth noticing that the performance of the R-ELM model increases with the number of samples K. This is explained by the fact that when K increases it is possible to obtain better estimates for H and Σ^H.

4 Conclusions

A variant of ELM that can be trained when the elements of the training set are random variables was proposed. The model was designed to be robust to the uncertainties of the training set. The resulting model, named R-ELM, is a regularized ELM without a regularization hyper-parameter.

The proposed model was used to predict cotton yarn quality measures collected from a real world textile industry and was compared to a linear regression model and a standard ELM. The results showed that the proposed model outperformed both benchmark methods in terms of ARMSE.

It was also possible to notice that the performance of proposed method is affected by the number of samples K that are used to calculate the matrix H. A possible topic for future research is to develop a method to define the optimum number of samples to be used.

References

[1] Aninda Ghosh and Pritam Chatterjee. Prediction of cotton yarn properties using support vector machine. *Fibers and Polymers*, 11:84–88, 2010.

[2] Khalid A. A. Abakar and Chongwen Yu. Application of genetic algorithm for feature selection in optimisation of SVMR model for prediction of yarn tenacity. *Fibers & Textiles in Eastern Europe*, 21:95–99, 2013.

[3] M. Selvanayaki, M. S. Vijaya, K. S. Jamuna, and S. Karpagavalli. An interactive tool for yarn strength prediction using support vector regression. In *proceedings of the Second International Conference on Machine Learning and Computing*, pages 335–339, Bangalore, February 2010.

[4] Khalid A. A. Abakar and Chongwen Yu. The performance of data mining techniques in prediction of yarn quality. *International Journal of Innovation, Management and Technology*, 4:390–392, 2013.

[5] You-Lo Hsieh, Xiao-Ping Hu, and Anjia Wang. Single fiber strength variations of developing cotton fibers - strength and structure of *G. hirsutum* and *G. barbedense*. *Textile Research Journal*, 70:682–690, 2000.

[6] Farzad Hosseinali. Investigation on the tensile properties of individual cotton (*Gossypium hirsutum L.*) fibers. Master's thesis, Texas Tech University, USA, August 2012.

[7] Sunil Kumar Sharma. Yarn hairiness questions and answers. *Spinning Textiles*, 7, 2013.

[8] Arindam Basu. Assessment of yarn hairiness. *Indian Journal Of Fiber And Textile Research*, 1999.

[9] G. B. Huang, Q. Y. Zuo, and C. K. Siew. Extreme learning machine: Theory and applications. *NeuroComputing*, 70:489–501, 2006.