Appropriate Software Techniques for Prediction Model to Attain Optimum Fabric Width of Single Jersey Finished Cotton Knitted Fabric in Textile Knitting Industries: A Detailed Survey

I BHUVANESHWARRI[1], A TAMILARASI[2]

[1] *Assistant Professor (Senior), Department of IT, Institute of Road and Transport Technology, Vasavi college post, Erode, INDIA*

[2] *Professor & Head, Department of Computer Applications, Kongu Engineering College, Perundurai, Erode, INDIA*

Abstract- ***Predicting knitted fabric properties and dimensional properties was a long-felt need for many years and a considerable amount of work was carried out using some simple, complex mathematical and empirical models. Fabric width of single jersey finished cotton knitted fabric is an important parameter that affects its comfort properties. This is also an important problem related to manufacturing costs. The most important problem in the knitting industry is attaining optimum fabric width during the garment manufacturing process. The textile industry needs to find and fix a flawless model to predict the best possible fabric width accurately to overcome the unwanted textile wastage and provide the most advantageous technology in knitted fabric manufacturing. This paper is a review that survey recent technologies applied for prediction models to attain optimum fabric width of single jersey finished cotton knitted fabric. This paper mainly focused on providing a brief review on the techniques such as feature selection using rough set theory, Data Mining (DM) techniques and managing of big data with Hadoop MapReduce framework. It summarizes the previous research works carried out on the above-mentioned techniques. Finally, it enlists the outcomes of the survey and gaps in this research work.***

Indexed Terms- ***Fabric width, rough set theory, data mining techniques, managing of big data***

I. INTRODUCTION

The objective of this research paper is to provide a survey on the software techniques such as feature selection using rough set theory, Data Mining (DM) techniques and managing of big data with Hadoop MapReduce framework. The detailed review of the basic concepts of various techniques is presented to provide an idea of the range of their potential usefulness to the textile knitting industries. In this paper, section II provides previous research work on rough set theory-based feature selection, section III discusses about previous research work done on prediction models, section IV includes studies on big data, and section V concludes the paper with its outcomes and gaps.

II. STUDIES ON ROUGH SET THEORY BASED FEATURE SELECTION METHODS

Today the world is driven by data. The presence of vagueness, incompleteness, and granularity in the generated data set gives an unreliable solution in the data analysis. Traditional approaches handled most of the problems in a crisp, deterministic and precise manner. But the real-world problems cannot be crisp and precise. Various methods have been developed by experts to address this issue. Now-a-days rough set theory based [1] model is used to handle vague, uncertain, imprecise and imperfect data. Table 1 gives the summary of research work on rough set theory-based feature selection carried out by various authors using various techniques.

Table.1: Summary of research work on rough set theory-based feature selection

References	Type of attributes / concept used	Application / data set
Golan & Ziarko (1995)[2]	Discrete attributes / Identify strong non-deterministic rules with estimates of decision probabilities with approximation space	Stock Market Analysis / 120 information vectors & 32 attributes
Hu *et al.* (2000) [3]	The neighborhood model is used to reduce numerical and categorical features by assigning different thresholds for different kinds of attributes / Positive region in rough set theory	Different data sets 32 attributes, 101-20000 samples
Zhang *et al.* (2005) [4]	Significance of features	Simple data set with 4 attributes , maximum of 300 tuples
Thangavel *et al.* (2005) [5]	Numeric and Non-numeric attributes/ Quick reduct method/ C4.5 Decision Tree algorithm	Medical/ data sets of UCI and the Human Immuno deficiency Virus data set 5 attributes, 8 data points
Thangavel *et al.* (2006) [6]	Categorical attributes / Used quick and variable precision rough set reduct algorithm to get reducts. Rule algorithm was used to generate rules from reducts	Human Immuno deficiency Virus data set 15 attributes,500 patient records
Wang *et al.* (2007) [7]	Discrete attributes / Genetic algorithm with PSO	Different data set from UCI repository, Maximum 6 attributes, 8124 Instances
Inbarani *et al.* (2007) [8]	Reduct concept	Web usage mining/ 9 attributes , 522 sessions
Fazayeli *et al.* (2008) [9]	Dimensionality reduction based on dependency of attributes	Simple data set from UCI repository 4 attributes with 6 objects
Chandra *et al.* (2016) [10]	Numerical & categorical attributes / Rough set, kNN, SVM	Automatic fabric Inspection System/ 400 fabric samples (training-200,test-200)

III. STUDIES ON PREDICTION MODELS

The various studies that have been carried out on the prediction model are outlined in this section. The high accuracy and automatic prediction models are required for textile industries to overcome several issues like time, cost and fabric wastages. A considerable amount of research works is carried out by the researchers to resolve the above-mentioned issues using naïve Bayesian classification, ANN and ANN with PSO methods. Table 2 presents a detailed summary of previous work done on prediction models using various techniques and their applications in textile industries.

Table.2: Summary of previous research work done on prediction models

References	Applications / Techniques used / No. of attributes or samples used	Remarks / descriptions
Golob *et al.* (2008) [11]	Prediction of combinations of dyes in textile printing paste formulations / NN / 1430 printed samples	The mean square error was 0.1 and more.
Behera & Goyal (2009) [12]	Performance parameters of airbag fabrics / ANN / 24 fabrics with 16 attributes	Achieved 12% of prediction error due to local minima and less number of samples

Murrells et al. (2009) [13]	Predicting spirality of single jersey 100% cotton fabric/ ANN / 60 samples	Achieved less MSE and high correlation coefficient with R of 0.976 and identified better parameters which affecting spirality
Al-Aidaroos *et al.*(2010) [14]	UCI machine learning repository / filter based feature selection & rough naïve Bayes (reduced error pruning)/ 2405 instances	Achieved 95% classification accuracy and suggested to improve accuracy more by better preprocessing method of data, attribute dependency calculations & identification of positive region
Furferi *et al.* (2012) [15]	Predicting characteristics of coated fabrics/ ANN / 90 (5 different fabrics with 18 different coating processes)	Effectively used to select process parameter with the maximum error ±10% with an average error of less than 5.5%
Bahlmann *et al.* (1999) [16]	Automated quality control of textile seams / NN - Self-organizing Feature Map based classification / 126 seam specimens	Lowering the manufacturing cost
Cichosz (2014) [17]	Data mining algorithms using R / Naïve bayes classifier	Explained naïve Bayes in DM using R tool
Akyol *et.al.* (2015) [18]	Prediction of drying time of wool / Filter based feature selection Reduced Error Pruning Tree based prediction model / 19 condition features and 1 target feature with 947 tuples (18,942 data points).	Provided better prediction with best feature subset with minimum error in minimum time
Jaouachi *et al.* (2015) [19]	Prediction of the sewing thread consumption of jean trousers / NN / 6 input parameters	Achieved R^2=0.973, provided accurate prediction model
Anupreet and Roy (2016) [20]	Prediction model for the single jersey cotton knitted fabric to predict shrinkage and fabric weight (g/m^2) / ANN/ 144 samples	MSE found to be lower for different reference states of fabric
Kumar (2016) [21]	Prediction of disease (positive or negative)/ ANN-PSO/ UCI repository medical data set/ maximum of 5000 tuples	Performance of ANN improved by the inclusion of the PSO

IV. STUDIES ON BIG DATA

This part of the review encompasses studies on big and massive data. The major challenge for the textile manufacturing industry these days is global competition and unforeseen demand instabilities. These factors force manufacturers to update and improve their production process to reduce their production costs and also to meet the demand deadlines. Table 3 presents a detailed summary of previous work done on managing of massive data using various techniques.

Table. 3: Summary of previous work done on scalable techniques

References	Application / Framework/ Techniques used	Description
Liu et al. (2005)[22]	large-scale text categorization/ SVM	Developed scalable system and investigated threshold tuning algorithms with respect to time complexity and classification accuracy

Wu et al. (2012)[23]	Hadoop MapReduce	Implemented ACO under MapReduce programming model but needed efficient problem partition strategy
Fernandez et al. (2014)[24]	Big data with cloud computing	Insight on computing environment, MapReduce and programming model
Chen and Zhang (2014)[25]	Survey paper on Big Data	Data-intensive applications, challenges, techniques and technologies on Big data

CONCLUSION

The following outcomes were made from the exhaustive survey. Although a great deal of work has been carried out on the application of various tools such as regression analysis, ANN, GA, the shortcomings have been identified. This present survey addresses these issues and provides useful information for the textile industry. Most of the research workers used naive Bayes classifications which are based on the independence of the attributes. But dependency among the attributes is also needed to identify redundant and irrelevant attributes in the data set efficiently. All the previous prediction models developed by the researchers are based on fewer samples. So, the scalable algorithms are needed for effective handling of the massive textile data set. Hadoop framework based scalable algorithms might be worked well both in time and space complexity to provide better performance and processing speed.The following gaps were pointed from the survey. So far, rough set theory-based feature selection is not applied in textile applications. Gaussian naïve Bayes classification-based prediction method using the Hadoop MapReduce framework is not yet carried out in textile applications. The growing amount of available data has become a serious challenge to data mining and machine learning techniques. The well-known classification methods that have been widely applied so far are no longer feasible in big data environments.

REFERENCES

[1] Pawlak, Z 1982, ‘Rough sets’, International journal of computer & information sciences, vol. 11, no. 5, pp. 341-356.

[2] Golan, RH & Ziarko, W 1995, ‘A methodology for stock market analysis utilizing rough set theory’, In Proceedings of 1995 Conference on Computational Intelligence for Financial Engineering (CIFEr), IEEE, pp. 32-40.

[3] Hu, K, Diao, L, Lu, Y & Shi, C 2000, ‘A heuristic optimal reduct algorithm’, Proceedings of International Conference on Intelligent Data Engineering and Automated Learning, Springer, Berlin, Heidelberg, pp. 139-144.

[4] Zhang, TF, Xiao, JM & Wang, XH 2005, ‘Algorithms of attribute relative reduction in rough set theory’, Dianzi Xuebao (Acta Electronica Sinica), vol. 33, no.11, pp. 2080-2083.

[5] Thangavel, KJ, Jaganathan, P, Pethalakshmi, A & Karnan, M 2005, ‘Effective classification with improved quick reduct for medical database using rough system’, BIME Journal, vol. 5, no. 1, pp. 7-14.

[6] Thangavel, K, Shen, Q & Pethalakshmi, A 2006, ‘Application of clustering for feature selection based on rough set theory approach’, AIML Journal, vol. 6, no. 1, pp. 19-27.

[7] Wang, X, Yang, J, Teng, X, Xia, W & Jensen, R 2007, ‘Feature selection based on rough sets and particle swarm optimization’, Pattern recognition letters, vol. 28, no. 4, pp. 459-471.

[8] Inbarani, HH, Thangavel, K & Pethalakshmi, A 2007, ‘Rough set based feature selection for web usage mining’, Proceedings of International Conference on Computational Intelligence and Multimedia Applications, IEEE, vol. 1, pp. 33-38.

[9] Fazayeli, F, Wang, L & Mandziuk, J 2008, ‘Feature selection based on the rough set theory and expectation-maximization clustering algorithm’, Proceedings of International Conference on Rough Sets and Current Trends in Computing, Springer, Berlin, Heidelberg, pp. 272-282.

[10] Chandra, JK, Majumdar, M & Sarkar, S 2016 ,

'Feature extraction and classification of woven fabric using optimized Haralick parameters: A rough set based approach', Proceedings of *2nd International Conference on Control, Instrumentation, Energy & Communication (CIEC)*, Kolkata, pp. 541-545.

[11] Golob, D, Osterman, DP & Zupan, J 2008, 'Determination of pigment combinations for textile printing using artificial neural networks', Fibres & Textiles in East Europe, vol. 16, pp. 93–98.

[12] Behera, BK & Goyal, Y 2009, 'Artificial neural network system for the design of airbag fabrics', Journal of Industrial Textiles, vol. 39, no. 1, pp. 45-55.

[13] Murrells, CM, Tao, XM & Xu, BG & Cheng, KP 2009, 'An artificial neural network model for the prediction of spirality of fully relaxed single jersey fabrics', Textile Research Journal, vol. 79, no. 3, pp. 227-234.

[14] Al-Aidaroos, K, Bakar, AA & Othman, Z 2010, 'Data classification using rough sets and naïve Bayes', Proceedings of International Conference on Rough Sets and Knowledge Technology, Springer, Berlin, Heidelberg, pp. 134-142.

[15] Furferi, R, Governi, L & Volpe, Y 2012, 'Modelling and simulation of an innovative fabric coating process using artificial neural networks', Textile Research Journal, vol. 82, no. 12, pp. 1282-1294.

[16] Bahlmann, C, Heidemann, G & Ritter, H 1999, 'Artificial neural networks for automated quality control of textile seams', Pattern recognition, vol. 32, no. 6, pp. 1049-1060.

[17] Cichosz, P 2014, 'Data mining algorithms: explained using R', John Wiley & Sons.

[18] Akyol, U, Tüfekci, P, Kahveci ,K & Cihan ,A 2015, 'A model for predicting drying time period of wool yarn bobbins using computational intelligence techniques', Textile Research Journal, vol. 85, no. 13, pp. 1367-1380.

[19] Jaouachi, B & Khedher, F 2015, 'Evaluation of sewed thread consumption of jean trousers using neural network and regression methods', Fibres & Textiles in Eastern Europe.

[20] Anupreet, K & Roy, K 2016, 'Prediction of shrinkage and fabric weight (g/m^2) of cotton single jersey knitted fabric using artificial neural network and comparison with general linear model', International Journal of Information Research and Review, vol. 03, no. 06, pp. 2541-2544.

[21] Kumar, GK 2016, 'An optimized particle swarm optimization-based ANN Model for clinical disease prediction', Indian Journal of Science and Technology, vol. 9, no. 21, pp. 1-7.

[22] Liu, TY, Yang, Y, Wan, H, Zeng, HJ, Chen, Z & Ma, WY 2005, 'Support vector machines classification with a very large-scale taxonomy', ACM SIGKDD Explorations Newsletter, vol. 7, no. 1, pp. 36-43.

[23] Wu, B, Wu, G & Yang, M 2012, 'A map reduce based ant colony optimization approach to combinatorial optimization problems', In Proceedings of Natural Computation (ICNC), Eighth International Conference, IEEE, pp. 728-732.

[24] Fernández, A, del Río, S, López, V, Bawakid, A, del Jesus, MJ, Benítez, JM & Herrera, F 2014, 'Big Data with Cloud Computing: an insight on the computing environment, MapReduce, and programming frameworks', Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 4, no. 5, pp.380-409.

[25] Chen, CP & Zhang, CY 2014, 'Data-intensive applications, challenges, techniques and technologies: A survey on Big Data', Information Sciences, vol. 275, no.3, pp. 14-47.