การพยากรณ์รอยพับสำหรับคุณภาพกระดาษในอุตสาหกรรมการผลิตกระดาษ

นางสาว เฟย เฟย หวัง

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาการคอมพิวเตอร์และสารสนเทศ ภาควิชาคณิตศาสตร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2551
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

CURL FORECASTING FOR PAPER QUALITY IN PAPERMAKING INDUSTRY

Ms. Feifei Wang

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science Program in Computer Science and Information

Department of Mathematics

Faculty of Science

Chulalongkorn University

Academic Year 2008

Copyright of Chulalongkorn University

510891

Thesis Title	CURL FORECASTING FOR PAPER QUALITY IN PAPERMAKING INDUSTRY
By	Feifei Wang
Field of Study	Computer Science and Information
Advisor	Siripun Sanguansintukul,Ph.D.
Co-Advisor	Professor Chidchanok Lursinsap,Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

S. Hannongbua ..... Dean of the Faculty of Science

(Professor Supot Hannongbua, Dr.rer.nat.)

THESIS COMMITTEE

Syphakant Phimoltares ..... Chairman

(Suphakant Phimoltares, Ph.D.)

Siripun Sanguansintukul ..... Advisor

(Siripun Sanguansintukul, Ph.D.)

C. Lursinsap ..... Co-Advisor

(Professor Chidchanok Lursinsap, Ph.D.)

P. Watanapongse ..... External Examiner

(Assistant Professor Pirawat Watanapongse, Ph.D.)

เฟย เฟย หวัง : การพยากรณ์รอยพับสำหรับคุณภาพกระดาษในอุตสาหกรรมการผลิตกระดาษ. (CURL FORECASTING FOR PAPER QUALITY IN PAPERMAKING INDUSTRY) อ.ที่ปรึกษาวิทยานิพนธ์หลัก : อ. ดร. สิริพันธุ์ สงวนสินธุกุล, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม: ศ. ดร. ชิดชนก เหลือสินทรัพย์, 70หน้า.

วิทยานิพนธ์นี้ได้นำเสนอโมเดลการพยากรณ์รอยพับสำหรับคุณภาพกระดาษ โดย Neural network พร้อมทั้งข้อมูลจากขั้นตอนการผลิตหลายแหล่งในการวิเคราะห์ของโมเดล ทั้งนี้การทดสอบคุณภาพของกระดาษและการควบคุมเป็นปัจจัยสำคัญของอุตสาหกรรมการผลิตกระดาษ ซึ่งส่งผลกระทบต่อขั้นตอนการผลิตโดยรวมและอนาคตของการตลาดกระดาษ เมื่อเปรียบเทียบดัชนีชี้วัดคุณภาพของกระดาษด้วยกันแล้ว ดัชนีการพับจะเป็นสิ่งที่กระทบกับการใช้งานของผู้บริโภคโดยตรง อีกทั้งยังเป็นสิ่งที่ยากในการทดสอบและควบคุมในขั้นตอนการผลิตจริงจำนวนมากจากฐานข้อมูลการผลิตซึ่งอาจส่งผลต่อคุณภาพของกระดาษในขั้นตอนสุดท้ายจะถูกทำให้อยู่ในรูปแบบที่เหมาะสมก่อนที่จะนำเข้าสู่ขั้นตอนการประมวลผล ซึ่งตัวโมเดลจะใช้หลักการของ MLP Neural network และทำการเปรียบเทียบผลระหว่างวิธีการของ Quasi-Newton และ Double Dogleg โดยใช้วิธี early stopping ในขั้นตอนสุดท้าย ได้เพิ่มการขั้นตอน Bootstrap เพื่อทำให้ผลการทดลองที่ได้จากโมเดลมีความถูกต้องมากยิ่งขึ้น

ภาควิชา คณิตศาสตร์ ลายมือชื่อนิสิต Fei-fei Wang

สาขาวิชา วิทยาการคอมพิวเตอร์และสารสนเทศ ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก Siripo Sangsinkul

ปีการศึกษา 2551 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม C.L.

5073633723 : MAJOR COMPUTER SCIENCE AND INFORMATION

KEYWORDS NEURAL NETWORK / PCA / STEP-WISE REGRESSION / NEWTON-BASED ALGORITHM

FEIFEI WANG: CURL FORECASTING FOR PAPER QUALITY IN PAPERMAKING INDUSTRY ADVISOR: SIRIPUN SANGUANSINTUKUL, Ph.D. CO-ADVISOR : PROF.CHIDCHANOK LUPSINSAP, Ph.D., 70pp.

This thesis presents a quality-forecasting model based on neural network for the paper making industry with different source data transaction processes. The paper quality test and control plays an essential role in the paper making industry, which affects the whole operation process and the future paper market. Compared with other paper quality indexes, paper curl is closer to terminal clients and more difficult to pretest and control in the actual working environment. Large-scale data from production database, which would potentially affect final paper quality, have been cleansed and abstracted. Modeling based on MLP neural network was designed to compare between Quasi-Newton algorithm and Double Dogleg with early stopping regularization in different source data sets. With bootstrap accuracy estimation, the final result has been evolved which would annotate the relationship between workflow data and paper curvature in a more constructive way.

Department : Mathematics Student's Signature : Feifei Wang

Field of Study : Computer Science and Information Technology Advisor's Signature : Siripun Sanguansintukul

Academic Year : 2008 Co-Advisor's Signature : C. Lu

Acknowledgments

I am deeply appreciated to my thesis advisors, Dr. Siripun Sanguansintukul and Professor Chidchanok Lursinsap, for their cherishing, helpful guidance and warmest encouragement. Without their keep-going support and care, this thesis could not be finished on time.

I am also appreciated to the thesis committee, Assistant Professor Pirawat Wattanipong from Kasetsart University and Dr. Suphakant Phimoltares, for their constructive suggestion and advice.

I would like to thank all my teachers for their kind supports and all my classmates and friends for their care during my study.

Contents

	Page
Abstract (Thai)	iv
Abstract (English)	v
Acknowledgements	vi
Contents	vii
List of Tables	ix
List of Figures	x
CHAPTER	
I INTRODUCTION	1
1.1 Motivation and Problem Description	1
1.2 The Objective of Research	1
1.3 The Scope of Study	2
II BACKGROUND, METHODLOGY AND MODEL DESIGN	4
2.1 Industry Background and Data Selection	4
2.1.1 Industry Background	4
2.1.2 Data Selection	5
2.2 MLP Neural Networks	9
2.2.1 Introduction to Neural Networks	9
2.2.2 MLP Neural Network Architecture	12
2.3 Learning Algorithms	14
2.3.1 Quasi-Newton Algorithm	14
2.3.2 Double Dogleg Algorithm	15
2.4 Dimension Reducing Methods	16
2.4.1 Principal Components Analysis	16
2.4.2 Step-wise Regression	18
2.5 Forecasting Process Flow Design	19
III EXPERIMENTAL APPLICATION	21
3.1 Data Physical Preprocessing	21

Chapter Page

3.2 Input Pruning 24
3.2.1 Data Normalization 24
3.2.2 Input Reduction 25
3.3 Tuning Neural Networks Models 27
3.3.1 Data Partition 27
3.3.2 Preliminary Training 28
3.3.3 Early Stop 30
3.3.4 Direct Connection 31
3.3.5 Number of Hidden Units 32
IV EXPERIMENT RESULT 34
4.1 Experimental Processing Results 34
4.2 Accuracy Evaluation 36
4.3 ROC Curve 40
V CONCLUSIONS AND FUTURE WORKS 44
5.1 Conclusions 44
5.2 Discussions 46
5.3 Future Works 47
REFERENCES 48
APPENDICES 50
Appendix A SAS Enterprise Miner and SEMMA Architecture 51
Appendix B Coding 55
VITAE 60

List of Tables

Table		Page
3.1	Related parameters description.....	23
3.2	Selected variables for second data set.....	26
4.1	Training results summaries.....	35
4.2	Test result analysis.....	39
4.3	Cross-tabular frequency table.....	40

List of Figures

Figure		Page
2.1	Stages involved in gloss paper production	5
2.2	Interface of sample source data	6
2.3	Source data selection flow	7
2.4	A neural network MLP architecture	14
2.5	Geometric interpretation of the first principal component	17
2.6	Forecasting process flow	20
3.1	Source data preprocessing analysis	22
3.2	Principal component analysis result	26
3.3	Multiple minima in error function	30
3.4	A skip-layer network	32
4.1	Actual curl vs. predicted curl in jumbo reel series	36
4.2	Actual curl vs. predicted curl	36
4.3	Increasing model accuracy design flow	37
4.4	Actual curl vs. predicted curl in jumbo reel series (bagging)	38
4.5	Actual curl vs. predicted curl (bagging)	38
4.6	One-Way frequencies for target value as "accept"	41
4.7	One-Way frequencies for target value as "reject"	42
4.8	Value of ROC areas	43
5.1	Detailed design layout	45
A.1	Enterprise Miner 4.3 Interface	52
A.2	SEMMA process illustration	53