APPLICATION OF SUPPORT VECTOR MACHINE AND NEURAL NETWORK MODELING IN THE PREDICTION OF CONCENTRATION OF DISPERSED PHASE OUTLET IN ROTATING DISC CONTACTOR (RDC) COLUMN

EZZATUL FARHAIN BINTI AZMI

A dissertation submitted in partial fulfillment of the

requirements for the award of the degree of

Master of Science (Mathematics)

Faculty of Science

Universiti Teknologi Malaysia

JANUARY 2013

To my beloved parents, Azmi bin Abdul Jalil and Rohiaza binti Murad. Also to all my families and supportive friends.

ACKNOWLEDGEMENT

First and foremost, all praise to be Allah, the Almighty, the Benevolent for His blessings in completing this project.

I would like to express my sincere appreciation to my supervisor, PM. Dr. Khairil Anuar bin Arshad for his guidance, knowledge, encouragement, critics and motivations in completing this report. Without his continued support, this thesis would not have been the same as presented here.

I wish to thank various people for their contribution to this project; Mr. Agos bin Arbain and Ms Fazlynda binti Murad, for their valuable technical support on this project; Muhammad Syazwan bin Ab Razak for his help in handling programming and computer; Farah Asyikin binti Ahmad, for her help in grammar checking; Nor Hamizah binti Miswan, Nurul Fariha binti Mokhter and Muhammad Rahimi bin Yaakob, for their advices and continuous supports; and anyone who helped me directly or indirectly in the completion of this project. Again, thank you very much.

I would also like to express my gratitude to all my lecturers in Department of Mathematics, Faculty of Science, UTM for every knowledge, guidance and supportive along my study life in UTM.

Finally, special thanks should be given to my family for their love, support and encouragement throughout my study. Your support really means a lot to me. I love you so much.

ABSTRACT

Liquid-liquid extraction is one of the most important separation processes that widely used in industries. Rotating Disc Contactor (RDC) column is one of the liquid-liquid extractor. Therefore, the study of liquid-liquid extraction in RDC column has become a very important subject to be discussed not just among the chemical engineers but mathematician as well. This project presents Support Vector Machine (SVM) and Neural Network modeling in the prediction of concentration of dispersed phase outlet in RDC column. SVM is an exciting Machine Learning technique that learns by example to sign labels to object and can be used for regression as well as classification purpose, while Neural Network is widely used as effective approach for handling nonlinear data especially in situations where the physical processes are not fully understood. Both modeling systems offer the potential for a more flexible and less error in forecasting. Thus, it can help to save time and reducing cost in conducting experiments. A Statistica software is utilized to help with the SVM modeling and a Matlab code is produced to run the Neural Network simulation in this project. The mean square error is calculated to compare the result between the two models. The analysis shows that both SVM and Neural Network modeling can predict the concentration of dispersed phase in RDC column but the SVM approach gives better result than the Neural Network approach.

ABSTRAK

Pengekstrakan cecair dengan cecair merupakan salah satu proses pemisahan paling penting yang digunakan secara meluas dalam banyak industri. Turus Pengekstrakan Cakera Berputar (RDC) adalah salah satu turus pengekstrakan cecair dengan cecair tersebut. Oleh itu, kajian pengekstrakan cecair dengan cecair dalam turus RDC telah menjadi subjek yang sangat penting yang akan dibincangkan bukan sahaja di kalangan jurutera kimia tetapi ahli matematik juga. Projek ini membentangkan model Sokongan Mesin Vektor (SVM) dan model Rangkaian Neural dalam meramal kepekatan fasa terserak dalam turus RDC. SVM merupakan suatu teknik pembelajaran mesin di mana ia belajar melalui teladan untuk menandatangani label kepada objek dan boleh digunakan untuk regresi serta tujuan pengelasan, manakala Rangkaian Neural digunakan secara meluas sebagai pendekatan yang efektif dalam mengendalikan data tidak linear terutamanya dalam situasi di mana proses fizikal tidak sepenuhnya difahami. Kedua-dua model SVM dan Rangkaian Neural menawarkan pendekatan yang lebih berpotensi dan ralat yang lebih kecil dalam sesuatu peramalan. Oleh itu, ia dapat membantu menjimatkan masa dan mengurangkan kos dalam menjalankan eksperimen. Satu perisian Statistica digunakan untuk membantu menyelesaikan model SVM dan kod Matlab dihasilkan untuk menjalankan simulasi Rangkaian Neural dalam projek ini. Purata jumlah perbezaan kuasa dua dihitung untuk membandingkan hasil antara kedua-dua model. Analisis menunjukkan bahawa kedua-dua model SVM dan Rangkaian Neural boleh meramal kepekatan fasa terserak dalam turus RDC tetapi pendekatan SVM memberikan hasil yang lebih baik daripada pendekatan Rangkaian Neural.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Liquid-liquid extraction is an important separation that widely used in industries. It was applied in chemicals, pharmaceuticals, refining, polymers and food industries. Specifically, chemicals use liquid-liquid extraction for water treatment, pharmaceuticals use for antibiotics production, refining use for aromatics, polymers use for Adiponitrile and food use for flavors. From these productions, we can see that liquid-liquid extraction can lead to a comfortable life for human. As example, production antibiotics in pharmaceutical can improve human healthiness.

Rotating Disc Contactor (RDC) columns are one of the extractors that used for this liquid-liquid extraction. It also has an extensive application in various industries. The performances of these columns indicate that they are more efficient and possess better operational flexibility. However, there is still some improving that researchers can do to enhance the performances.

Recently, many researches on forecasting have been conducted. This is because prediction result can lead to good decision making and planning. For example,

forecasting in production of liquid-liquid extraction can lead to precaution and some improving action in the extractor to give a better result. It also could cut the cost, time and energy in conducting the extraction. In order to forecast, academicians can help by constructing a model and understanding the applications and problems so that some actions and improvement could be taken.

Modeling is an extremely powerful tool, a framework for research, debate and planning, which provides a valuable source of information for decision making. Arnold Neumaier (2004) said that mathematical modeling is the art of translating problems from an application area into tractable mathematical formulations whose theoretical and numerical analysis provides insight, answers, and guidance useful for the originating application. Mathematical modeling is a bridge between the study of mathematics and the applications of mathematics to various fields. It is able to give more understanding of the system modeled and gives precision and direction for the problem solution. It is impossible to imagine modern science without the wide application of mathematical modeling. Therefore, this research is about applying the mathematical modeling using Support Vector Machine (SVM) and Neural Network for liquid-liquid extraction in RDC column.

1.2 Background of Problem

Liquid-liquid extraction is an extraction or separation of a substance from one liquid phase into another liquid phase. Separation is achieved by adding a liquid solvent phase to the original liquid carrying the components to be extracted. One of the phases must be dispersed into droplets in the other, continuous phase to achieve a sufficiently large mass-transfer interface. This extraction is widely used in the industries such as in chemical, pharmaceuticals, petroleum and food industry.

Difference condition and needs in liquid-liquid extraction to produce what we want in the results makes it need difference system to extract or dissolve between the bonds. That is why there are many types column been created and used for liquid-liquid extraction such as Packed column, Scheibel column, Karr column, Pulsed column and Rotating Disc Contactor (RDC) column. In this project, we study about the liquid-liquid extraction in RDC column.

Liquid-liquid extraction in RDC column involve process that the independent variables which are the inputs such as rotor speed, dispersed phase flow rate, concentration of continuous inlet and concentration of dispersed inlet makes the extraction for the results of the output which is the concentration of dispersed phase outlet. However, we do not know what actually happen in the RDC column because of the complexity process. Indeed, we need the end result.

At the beginning, the researches use mathematical simulation to get the end result (the output). In the study of Maan (2005), she use reverse modeling for this liquid-liquid extraction in the RDC column. There also other studies in this area involving prediction of the output. However, there are still some ways we can improve the prediction. Now, we want to improve the modeling by the new method. Therefore, in this project, we introduce Support Vector Machine and Neural Network modeling to predict the output (concentration of dispersed phase outlet). We also compare the results from both of the modeling to see the effectiveness. Hopefully this project can help the researches to improve the process and build the better RDC column.

1.3 Rational of Study

This study is subjected to a few of rational. Firstly, liquid-liquid extraction is very important and needed in the industries. It used for industries' productions such as water treatment in chemical industry and lube oil quality improvement in petroleum

industry. Besides, we do not know the exact process happen in the RDC column but the result of the process is very significant. Therefore, we would like to study and develop SVM and Neural Network modeling in this area as these are two of the popular methods that are being utilized in recent years. We concern in predicting the concentration of dispersed phase outlet in RDC column. We can determine the output without doing the experiment and predict thousands of data. The prediction will offer to help the researchers and engineers in improving the RDC column for a better result.

1.4 Problem of Statement

The prediction of concentration of dispersed outlet in RDC column needs to be improved as it indicate the effectiveness of mass transfer and drop breakage liquid-liquid extraction in RDC column. This project tries to answer some questions such as can SVM and Neural Network modeling be applied to help the engineers? Based on the data obtained from the experiment, how do we apply the modeling to get the prediction? Which type of modeling technique is the most suitable to describe the concentration of dispersed outlet in RDC column without the use of time consuming activities?

1.5 Objectives of the Study

The objectives of this study are:

1. To study the application of Support Vector Machine and Neural Network in the field of chemical engineering.

2. To apply a mathematical model using SVM and Neural Network for prediction of concentration of dispersed phase outlet in RDC column.
3. To compare the actual output and predicted output derived from the SVM and Neural Network.

1.6 Scope of Study

The study is focusing on SVM and Neural Network modeling for predicting the concentration of dispersed phase outlet in RDC column. The data is acquired from the researchers at the University of Bradford under contract to Separation Processes Service, AEA Technology, Harwell. There are 256 sets of data available and this data set consist of four columns of input data which are rotor speed, dispersed phase flow rate, concentration of continuous inlet, concentration of dispersed inlet and one columns of output data which is the concentration of dispersed phase outlet. STATISTICA and MATLAB software are used to develop the method in the model. Both methods will also be implemented as comparison to see their effectiveness.

1.7 Importance of the Study

This study is carried out with the main objective of evaluating SVM and Neural Network modeling. Based on the results, it is hoped that this study will be able to:

1. Develop the understanding on SVM and Neural Network modeling.
2. Provide basis for researchers on how to apply mathematical modeling using SVM and Neural Network in chemical engineering research.

3. Improve and encourage works in exploring the advantages of SVM, Neural Network methods and liquid-liquid extraction in RDC column.
4. Improve the liquid-liquid extraction process and build the better RDC column.

1.8 Thesis Outline

The proposal begins with Chapter 1 which consists of the introduction, background of problem, rational of study, problem statement, objectives, scope and importance of the study. This is followed by Chapter 2 that discusses about the literature review on the liquid-liquid extraction, RDC column, SVM and Neural Network modeling. The research planning, operational framework and research methodology are discussed in Chapter 3. Next, Chapter 4 represents the result and analysis of this study. Finally, from the result and analysis, conclusion can be made, some contribution mentioned and several recommendations are stated in Chapter 5.

REFERENCES

Alpaydin, & Ethem. (2004). Introduction to Machine Learning. Retrieved October 15, 2012, from http://books.google.com.

Boolchandani, D., & Sahula, V. (2011). Exploring Efficient Kernel Functions for Support Vector Machine Based Feasibility Models for Analog Circuits. *International Journal of Design, Analysis and Tools for Circuits and Systems.* 1(1): 1-10.

Cao, L.J., & Francis, E. H. (2003). Support Vector Machine With Adaptive Parameters In Financial Time Series Forecasting. *IEEE Transactions On Neural Networks*. 14(6): 1506-1518.

Chen, B.J., Chang,M.W., & Lin C.J. (2004). Load Forecasting Using Support Vector Machines: A Study on EUNITE Competition 2001. *IEEE Transaction on Power Systems*. 19(4): 1821-1830.

Dorofki, M., Elshafie, A.H. Jaafar, O. & et.al. (2012). Comparison of Artificial Neural Network Transfer Functions Abilities to Simulate Extreme Runoff Data. *2012 International Conference on Environment, Energy and Biotechnology*. 33(2012): 39-44.

Fi, J.C., & Yen, C.C. (2001). A Counterpropagation Fuzzy-Neural Network Modeling Approach To Real Time Streamflow Prediction. *Journal of Hydrology*. 245(2001): 153-164.

Gordon, G. (2003). *Support Vector Machines and Kernel Methods*. Pittsburgh.

Han, D., Chan, L., & Zhu, N. (2007). Flood Forecasting Using Support Vector Machines. *Journal of Hydroinformatics.* 9(4): 267-276. IWA.

Hule, P.V., Murthy, B.N.,& Mathpati, C.S. (2011). Computational Fluid Dynamics of Two- Opposed-Jet Microextractor. *International Journal of Chemical Engineering.* 2012(2012): 1-11.

Junaida binti Sulaiman (2005). *Experimenting The Applicability Of Support Vector Machine Regression In Rainfall Prediction.* Master Thesis, Universiti Teknologi Malaysia.

Kaastra, I., & Boyd, M. (1995). Designing A Neural Network For Forecasting Financial And Economic Time Series. *Journal of Neurocomputing*. 10(1996): 215-236. Elsevier.

Luk, K.C., Ball, J.E., & Sharma A. (2001). An Application Of Artificial Neural Networks For Rainfall Forecasting. *Journal of Mathematical and Computer Modelling*. 33(2001): 883-699. Elsevier.

Maan, N. (2005). *Mathematical Modeling of Mass Transfer in Multi-stage Rotating Disc Contactor Column.* Doctor Philosophy, Universiti Teknologi Malaysia.

Nawi, N.H. (2010). *Model of Mass Transfer of Extraction in Rotating Disc Contactor Column.* Master Thesis, Universiti Teknologi Malaysia.

Noble, W.S. (2006). What is Support Vector Machine?. *Nature Biotechnology*. 24(12): 1565- 1567. Nature Publishing Group.

Neumaier, A. (2004). Modeling Languages in Mathematical Optimization. In J. Kallrath (Ed.*), Applied Optimization.* Kluwer: Boston.

Pao, S.Y., Shien, T.C, & Chang, I.F. (2005). Support Vector Regression For Real-Time Flood Stage Forecasting. *Journal of Hydrology*. 328: 704-716. Elsevier.

Rahayu binti Hamid (2004). *Bioactivity Prediction and Compound Classification Using Neural Network and Support Vector Machine: A Comparison*. Master Thesis, Universiti Teknologi Malaysia.

Rao M.A., & Srinivas J. (2003). *Neural Networks Algorithms and Applications*. United Kingdom: Alpha Science.

Rich, E., & Knight, K. (1996). *Artificial Intelligence Series*. McGraw-Hill.

Sansom, D.C., Downs, T., & Saha, T.K. (2003). Evaluation Of Support Vector Machine Based Forecasting Tool In Electricity Price Forecasting For Australian National Electricity Market Participants. *Journal of Electrical & Electronics Engineering, Australia.* 22(3): 227-233.

Schalkoff, R. (1997). *Artificial Neural Networks. Overview: Artificial Neural Networks and Neural Computing*, USA: 1:1-3.

Taylor J.Y., & Buizza R. (2002). Neural Network Load Forecasting With Weather Ensemble Predictions. *IEEE Transactions On Power Systems*. 17(3): 626-632.

Wan Nurul Aiffah binti Ismail (2010*). Modeling and Performance Analysis of Rotating Disc Contactor (RDC) Column.* Master Thesis, Universiti Teknologi Malaysia.

Wei, H., Nakamori, Y., & Wang S.Y. (2004). Forecasting Stock Market Movement Direction With Support Vector Machine. *Computers & Operations Research.* 32 (2005): 2513-2522. Elsevier.

Zhang G.P., & Qi M. (2005). Neural Network Forecasting For Seasonal And Trend Time Series. *European Journal of Operational Research*. 160(2): 501-514.