

Non-Invasive Forecast for Various Diseases

by

Jian Gong

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science in Mathematics

Faculty of Science and Technology

University of Macau

2011

Approved by _____

Supervisor

Date _____

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at the University of Macau, I agree that the Library and the Faculty of Science and Technology shall make its copies freely available for inspection. However, reproduction of this thesis for any purpose or by any means shall not be allowed without my written permission. Authorization is sought by contacting the author at

Address: Faculty of Science and Technology,
University of Macau,
Av. Padre Tomas Pereira, S.J. Taipa,
Macao, China.

Telephone: (853)62910651

E-Mail: fen_hong_nan_hai@163.com



Signature _____

Date _____

Non-Invasive Forecast for Various Diseases

by Jian Gong

Thesis Supervisor:

Professor Tao Qian

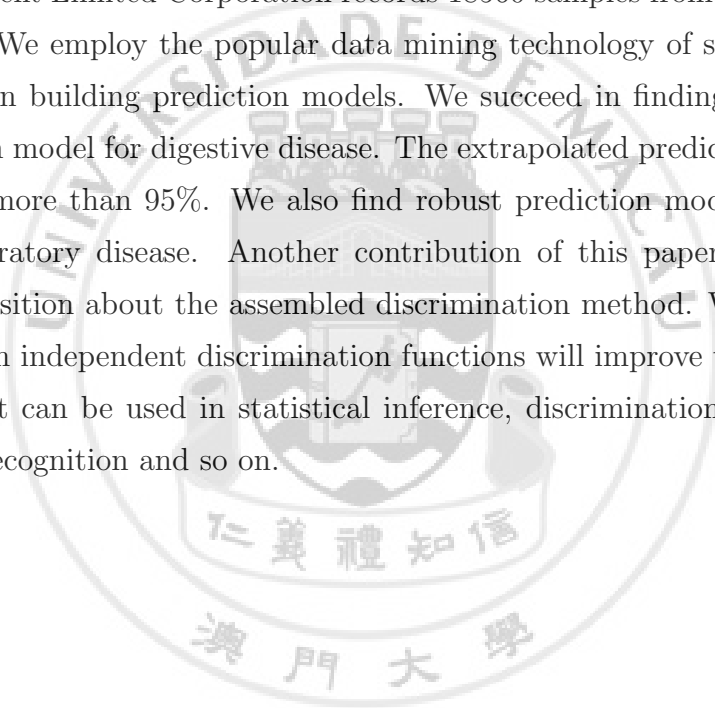
Department of Mathematics

University of Macau



Abstract

In this paper non-invasive disease prediction models are created for stroke, digestive disease and respiratory disease. The invasive disease prediction which causes wounds to human bodies is unsatisfactory. This is the primary motivation for our research. In our experimentation design three kinds of factors including meridian indices, past disease records and daily living habits are supposed to be potential indicators for determination of non-invasive disease forecast. The database provided by Hong Kong Intelligent Health Management Limited Corporation records 18560 samples from the year 2006 to 2010. We employ the popular data mining technology of support vector machine in building prediction models. We succeed in finding high-quality prediction model for digestive disease. The extrapolated prediction accuracy achieves more than 95%. We also find robust prediction models for stroke and respiratory disease. Another contribution of this paper results from the proposition about the assembled discrimination method. We prove that increase in independent discrimination functions will improve the prediction results. It can be used in statistical inference, discrimination analysis and pattern recognition and so on.



Contents

1	Introduction	1
2	Biochemistry Basis for Meridian Index	4
3	Summary of Support Vector Machine	6
3.1	Basic Models of SVM	6
3.2	Nonlinear Separable Generalization of SVM	9
3.2.1	Dual Transformation	10
3.2.2	Nonlinear Generalization and Kernel Functions	11
4	Research Design	13
5	Modeling and Forecast	15
5.1	Data Analysis	15
5.2	Assembled Predicting Method	18
5.3	Prediction for Stroke Disease	22

5.4	Prediction for Digestive Disease	24
5.5	Prediction for Respiratory Disease	26
5.6	Result Analysis	28
6	Conclusions	31
	Bibliography	33



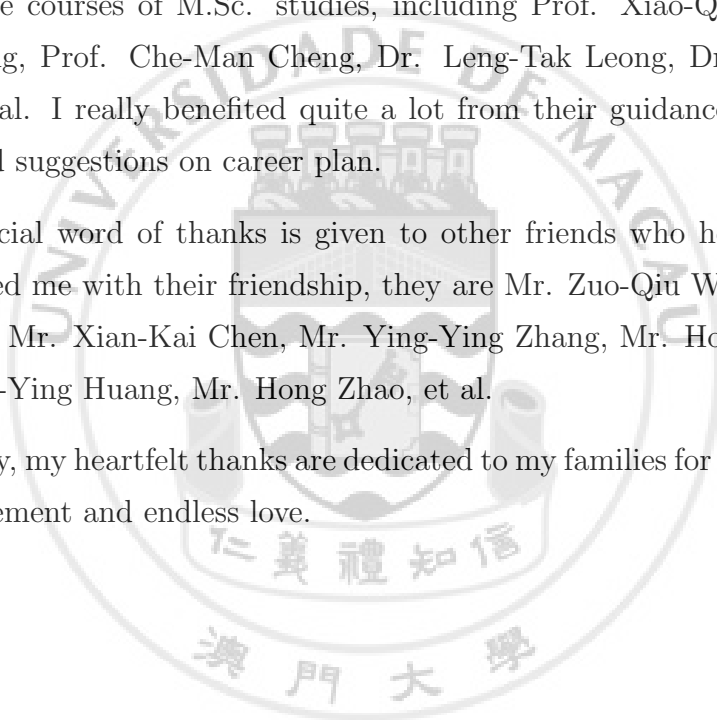
ACKNOWLEDGEMENT

First and foremost, I wish to express my sincerest and deepest gratitude to my supervisor, Prof. Tao Qian, for his inspiring guidance, support and constant encouragement throughout the whole period of my M.Sc. studies. His valuable advice and patience help me overcome a great many difficulties and doubts during the preparation of this thesis.

I would also deeply acknowledge other professors who have taught me during the courses of M.Sc. studies, including Prof. Xiao-Qing Jin, Prof. Deng Ding, Prof. Che-Man Cheng, Dr. Leng-Tak Leong, Dr. Seak-Weng Vong, et al. I really benefited quite a lot from their guidance in academic study and suggestions on career plan.

A special word of thanks is given to other friends who helped me and encouraged me with their friendship, they are Mr. Zuo-Qiu Weng, Mr. Zhi-Xiong Li, Mr. Xian-Kai Chen, Mr. Ying-Ying Zhang, Mr. Hong-Kui Pang, Mr. Ning-Ying Huang, Mr. Hong Zhao, et al.

Finally, my heartfelt thanks are dedicated to my families for their constant encouragement and endless love.



DECLARATION

The author declares that this thesis represents his own work with Professor Tao Qian, the author's supervisor. All the work is done under the supervision of Professor Qian during the period 2008-2010 for the degree of Master of Science in Mathematics at the University of Macau. The results in this thesis, unless otherwise stated or indicated, have not been previously included in any thesis, dissertation or report submitted to any institution for a degree, diploma or other qualification, or for publication by the author, and to the author's knowledge, by anyone else.



Jian Gong