UNIVERSITI PUTRA MALAYSIA

HYBRID PERFORMANCE MEASURES AND MIXED EVALUATION METHOD FOR DATA CLASSIFICATION PROBLEMS

MOHAMMAD BIN HOSSIN

FSKTM 2012 22

HYBRID PERFORMANCE MEASURES AND MIXED EVALUATION METHOD FOR DATA CLASSIFICATION PROBLEMS

By

MOHAMMAD BIN HOSSIN

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfillment of the Requirements for the Degree of Doctor of Philosophy.

April 2012

DEDICATION

This thesis is dedicated to:

My lovely mother Yah Bt. Tahir,

My late beloved father Hossin B. Mat, and

My brothers Huslan, Jamali and Eric

Abstract of this thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirement for the degree of Doctor of Philosophy

HYBRID PERFORMANCE MEASURES AND MIXED EVALUATION METHOD FOR DATA CLASSIFICATION PROBLEMS

By

MOHAMMAD BIN HOSSIN

April 2012

Chairman : Associate Professor Dr. Md. Nasir Sulaiman, PhD

Faculty : Computer Science and Information Technology

This study investigates two different issues of performance measure in data classification problem. First, this study examines the use of accuracy measure as a discriminator for building an optimized Prototype Selection (PS) algorithm. Second, this study evaluates the current evaluation practices for evaluating and comparing the two performance measures.

From the literature, the use of accuracy could lead to the underperforming of the evaluation process due to less distinctive and less discriminable values, and also unable to perform optimally when confronted with imbalanced class problem. Interestingly, the accuracy measure is still widely used in evaluating data classification problem. On the evaluation analysis, many previous studies emphasize on the generalization ability in evaluating and comparing the performance measures. Only few efforts have been dedicated to evaluate and compare the performance measures using different performance characteristics. In fact, no previous studies

employ mixed evaluation method in evaluating and comparing the performance measures.

For tackling the first issue, this study has successfully proposed several hybrid measures through the combination of accuracy with precision and recall measures. These hybrid measures are known as Optimized Accuracy with Conventional Recall-Precision (OACRP) and Optimized Accuracy with Extended Recall-Precision version 1 and version 2 (OAERP1 and OAERP2). More importantly, the OAERP1 and OAERP2 measure have been extended for evaluating multi-class problem. For the second issue, this study has proposed mixed evaluation method to evaluate the performance of two performance measures through different performance characteristics.

For a systematic analysis, the mixed evaluation method is implemented into two stages. First, the hybrid measures are compared and analyzed against the accuracy measure based on their produced-values through different classification problems with different class distribution problems. Second, the hybrid measures are compared and analyzed empirically against the accuracy measure and other selected performance measures based on generalization ability using three selected PS algorithms (MCS, LVQ21 and GA) and large benchmark datasets.

In the first evaluation stage, the OAERP2 measure has shown better produced-value against accuracy, OACRP and OAERP1 measures in terms of distinctiveness, discriminability, informativeness, favors towards minority class, and degree of consistency and discriminatory. In the second evaluation stage, almost all selected

algorithms that optimized by OAERP2 measure are able to produce better generalization ability against its original measure and other selected performance measures. Moreover, the GA model that was optimized by OAERP2 measure (GA_{oe2}) performed significantly and statistically differently as compared to other OAERP2-based models through win-draw-loss evaluation method and two non-parametric tests. Interestingly, the GA_{oe2} model also performed significantly and statistically differently as compared to nine additional PS algorithms in terms of testing error and storage requirements.

From all evaluations, it clearly reveals that the OAERP2 measure is able to choose a better solution during the classification training. As a result, it leads towards a better trained PS classifier with better generalization ability. On the other hand, the mixed evaluation method has enabled this study to evaluate and compare the studied performance measures systematically and comprehensively via different performance characteristics.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

PENGUKUR PRESTASI HIBRID DAN KAEDAH PENILAIAN CAMPURAN UNTUK PERMASALAHAN KLASIFIKASI DATA

Oleh

MOHAMMAD BIN HOSSIN

April 2012

Pengerusi : Profesor Madya Dr. Md. Nasir Sulaiman, PhD

Fakulti : Sains Komputer dan Teknologi Maklumat

Kajian ini mengkaji dua isu berbeza tentang pengukur prestasi bagi permasalahan klasifikasi data. Pertama, kajian ini meneliti penggunaan pengukur ketepatan sebagai diskriminator untuk membina algorithma Seleksi Prototaip (SP) yang optimum. Kedua, kajian ini juga mengkaji praktis penilaian yang terkini bagi menilai dan membandingkan dua pengukur prestasi.

Dalam kajian lepas, penggunaan ketepatan boleh menyebabkan proses penilaian di bawah tahap pencapaian disebabkan oleh nilai kurang unik dan kurang daya boleh-beza, serta tidak boleh bertindak secara optimum apabila berhadapan dengan permasalahan kelas tak-seimbang. Menariknya, pengukur ketepatan masih lagi digunakan secara meluas dalam menilai permasalahan klasifikasi data. Disudut analisis penilaian, kebanyakkan kajian lepas menekankan kebolehan pengitlakan dalam menilai dan membandingkan pengukur prestasi. Didapati hanya sedikit kajian yang dijalankan untuk menilai dan membandingkan pengukur prestasi menggunakan cirian prestasi yang berbeza. Malah, tiada kajian lepas menggunakan kaedah penilaian campuran dalam menilai dan membandingkan pengukur prestasi.

Untuk menyelesaikan isu pertama, kajian ini telah mencadangkan beberapa pengukur hibrid melalui kombinasi pengukur kejituan dan *precision* dan *recall*. Pengukur-pengukur hibrid ini dikenali sebagai *Optimized Accuracy with Conventional Recall-Precision* (OACRP) dan *Optimized Accuracy with Extended Recall-Precision* versi 1 dan 2 (OAERP1 dan OAERP2). Yang lebih penting, pengukur OAERP1 dan OAERP2 telah dikembangkan untuk menilai permasalahan multi-kelas. Untuk isu kedua, kajian ini telah mencadangkan kaedah penilaian campuran untuk menilai prestasi dua pengukur melalui cirian prestasi yang berbeza.

Untuk analisis yang sistematik, kaedah penilaian campuran ini dilaksanakan dalam dua peringkat. Pertama, pengukur hibrid dibandingkan dan dianalisis secara perbandingan dengan pengukur ketepatan berdasarkan nilai-hasil melalui permasalahan klasifikasi yang berbeza serta permasalahan distribusi kelas. Kedua, pengukur hibrid ini seterusnya dibandingkan dan dianalisis dengan pengukur ketepatan dan beberapa pengukur prestasi terpilih secara empirikal berdasarkan kebolehan pengitlakan melalui tiga algoritma terpilih (MCS, LVQ21 dan GA) serta set data tanda aras yang banyak.

Dalam penilaian peringkat pertama, pengukur OAERP2 telah menunjukkan nilai-hasil yang lebih baik berbanding pengukur ketepatan, OACRP dan OAERP1 berdasarkan keunikan, kebolehbezaan, daya maklumat, bantuan ke arah kelas minoriti, dan darjah ketekalan dan kebolehbezaan. Untuk penilaian peringkat kedua, hampir keseluruhan algoritma terpilih yang dioptimumkan oleh pengukur OAERP2 menghasilkan kebolehan pengitlakan yang lebih baik berbanding pengukur asal dan beberapa pengukur prestasi terpilih yang lain. Selain itu, model GA yang

dioptimumkan oleh pengukur OAERP2 (GA_{oe2}) menunjukkan prestasi yang signifikan dan perbezaan signifikan secara statistik berbanding dengan model lain yang berasaskan OAERP2 melalui kaedah penilaian menang-seri-kalah dan dua ujian bukan parametrik. Yang lebih menarik, model GA_{oe2} ini juga menunjukkan prestasi yang signifikan dan perbezaan yang signifikan secara statistik berbanding sembilan algoritma SP tambahan berdasarkan nilai ralat dan keperluan penyimpanan.

Dari semua penilaian, ini jelas menunjukkan bahawa pengukur OAERP2 mampu memilih solusi yang lebih baik semasa latihan klasifikasi. Hasilnya, ia memimpin ke arah pengelas SP terlatih yang lebih baik dengan kebolehan pengitlakan yang baik. Selain itu, melalui kaedah penilaian campuran telah membolehkan kajian ini menilai dan membandingkan pengukur prestasi yang diuji secara sistematik dan menyeluruh melalui cirian prestasi yang berbeza.

ACKNOWLEDGEMENTS

Praise to Allah and our beloved Prophet Muhammad (PBUH).

I would like to heartily express my deepest indebtedness and thankfulness to my supervisor Associate Professor Dr. Hj. Md Nasir Sulaiman, who well-guided me to complete my doctoral study at Universiti Putra Malaysia. I am also highly thankful to my supervisory committees Associate Professor Dr. Norwati Mustapha and Associate Professor Dr. Rahmita Wirza Rahmat for their incredible help, comment and sharing experience to improve my doctoral research. I would like to acknowledge Dr. Aida Mustapha for her intellectual discussion and advised in writing my journal papers.

My full gratitude also goes to my lovely mother for educating me and infinite support to finish my study. To all my brothers, thanks for your great support and understanding. Also special thanks to all my friends at UPM and UNIMAS especially for their great support, comment and beneficial discussion during my study period.

Finally, I also express my special appreciation to Universiti Malaysia Sarawak and Minister of Higher Education, Malaysia for giving me an opportunity and scholarship to further my doctoral study.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Md. Nasir Sulaiman, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Norwati Mustapha, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Rahmita Wirza O.K Rahmat, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

_____

BUJANG BIN KIM HUAT, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

DECLARATION

I declare that the thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently, submitted for any other degree at Universiti Putra Malaysia or at any other institution.

_____

MOHAMMAD BIN HOSSIN

Date: 27 April 2012

TABLE OF CONTENTS