FUZZY NEURAL NETWORKS WITH GENETIC ALGORITHM-BASED LEARNING METHOD

M. REZA MASHINCHI

UNIVERSITI TEKNOLOGI MALAYSIA

FUZZY NEURAL NETWORKS WITH GENETIC ALGORITHM-BASED LEARNING METHOD

M. REZA MASHINCHI

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy (Computer Science)

Faculty of Computing
Universiti Teknologi Malaysia

MARCH 2013

To my father, and family members

ACKNOWLEDGEMENT

I am deeply grateful to my supervisor, Prof. Dr. Ali Selamat, for his detailed and constructive comments and for his understanding and important support throughout this work. Also, my sincere thanks are due to the official examiners Assoc. Prof. Dr. Muhammad Suzuri Hitam, and, Prof. Dr. Siti Mariyam Shamsuddin for their detailed review, constructive criticism and excellent advice during the preparation of this thesis. I would also like to thank my father and brother for their constructive academic sharing.

ABSTRACT

This thesis is on the reasoning of artificial neural networks based on granules for both crisp and uncertain data. However, understanding the data in this way is difficult when the data is so complex. Reducing the complexity of the problems that these networks are attempting to learn as well as decreasing the cost of the learning processes are desired for a better prediction. A suitable prediction in artificial neural networks depends on an in-depth understanding of data and fine tracking of relations between data points. Inaccuracies of the prediction are caused by complexity of data set and the complexity is caused by uncertainty and quantity of data. Uncertainties can be represented in granules, and the reasoning based on granules is known as granular computing. This thesis proposed an improvement of granular neural networks to reach an outcome from uncertain and crisp data. Two methods based on genetic algorithms (GAs) are proposed. Firstly, GA-based fuzzy granular neural networks are improved by GA-based fuzzy artificial neural networks. They consist of two parts: granulation using fuzzy c-mean clustering (FCM), and reasoning by GA-based fuzzy artificial neural networks. In order to extract granular rules, a granulation method is proposed. The method has three stages: construction of all possible granular rules, pruning the repetition, and crossing out granular rules. Secondly, the two-phase GA-based fuzzy artificial neural networks are improved by GA-based fuzzy artificial neural networks. They are designed in two phases. In this case, the improvement is based on alpha cuts of fuzzy weight in the network connections. In the first phase, the optimal values of alpha cuts zero and one are obtained to define the place of a fuzzy weight for a network connection. Then, in the second phase, the optimal values of middle alpha cuts are obtained to define the shape of a fuzzy weight. The experiments for the two improved networks are performed in terms of generated error and execution time. The results tested were based on available rule/data sets in University of California Irvine (UCI) machine learning repository. Data sets were used for GA-based fuzzy granular neural networks, and rule sets were used for GA-based fuzzy artificial neural networks. The rule sets used were customer satisfaction, uranium, and the datasets used were wine, iris, servo, concrete compressive strength, and uranium. The results for the two-phase networks revealed the improvements of these methods over the conventional one-phase networks. The two-phase GA-based fuzzy artificial neural networks improved 35% and 98% for execution time, and 27% and 26% for the generated error. The results for GA-based granular neural networks were revealed in comparison with GA-based crisp artificial neural networks. The comparison with other related granular computing methods were done using the iris benchmark data set. The results for these networks showed an average performance of 82.1%. The results from the proposed methods were analyzed in terms of statistical measurements for rule strengths and classifier performance using benchmark medical datasets. Therefore, this thesis has shown GA-based fuzzy granular neural networks, and GA-based fuzzy artificial neural networks are capable of reasoning based on granules for both crisp and uncertain data in artificial neural networks.

ABSTRAK

Tesis ini menyelidik taakulan bagi rangkaian neural buatan berdasarkan granul untuk kedua-dua data jelas dan tidak jelas. Kaedah pemahaman data melalui cara ini adalah sukar apabila kandungan data adalah kompleks.Untuk mengurangkan kekompleksan masalah yang cuba dipelajari oleh rangkaian ini dan juga mengurangkan kos proses pembelajarannya, teknik ramalan yang lebih baik adalah diperlukan. Ramalan yang sesuai dalam rangkaian neural buatan bergantung kepada kebolehan untuk memahami isi kandungan data dengan mendalam dan juga kebolehan untuk mengenal pasti hubungan antara data. Ketakpastian dan kepelbagaian jenis data juga akan menjadikan hasil ramalan yang tidak tepat. Ketakpastian terhadap jenis data disebabkan oleh kekompleksan jenis data tersebut dan juga set data yang mengandungi tahap ketakpastian yang kompleks. Ketakpastian boleh diwakili dengan granul dan taakulan yang dikenali sebagai pengkomputeran granular. Tesis ini menggunakan rangkaian neural granular untuk mencapai hasil daripada data yang tidak pasti dan jelas. Dua kaedah telah diperkembangkan berdasarkan algoritma genetik. Rangkaian granular kabur berasaskan algoritma genetik (GA) telah diperkembangkan menggunakan rangkaian neural buatan kabur berasaskan GA. Rangkaian neural granular kabur berasaskan GA mengandungi dua bahagian: granulasi menggunakan pengelompokan min-c kabur, dan taakulan oleh rangkaian neural buatan kabur berasaskan GA. Untuk mengekstrak peraturan granular kaedah granulasi yang diterokai mengandungi tiga peringkat, iaitu pembinaan semua peraturan granular yang mungkin, pemangkasan data yang berulang dalam set data dan pengurangan peraturan granular yang telah digunakan. Rangkaian neural buatan kabur berasaskan GA berfasa dua telah direka bentuk dalam dua fasa. Dalam keadaan ini peningkatannya berdasarkan kepada nilai *alfa-cut* dalam rangkaian neural. Dalam fasa pertama nilai optimum *alfa-cut* adalah sifar dan boleh diperoleh bagi menentukan pemberat set kabur kepada rangkaian neural. Dalam fasa kedua nilai optimum untuk alfa-cut tengah diperoleh untuk menentukan bentuk set kabur. Uji kaji untuk dua rangkaian neural yang telah dipertingkatkan telah dijalankan berdasarkan kepada jumlah ralat yang dihasilkan dan masa yang diambil bagi melaksanakan uji kaji tersebut. Hasil uji kaji berdasarkan set data mesin pembelajaran repositori di University of California Irvine (UCI). Set data yang digunakan untuk rangkaian neural granular kabur berasaskan GA dan set peraturan digunakan untuk rangkaian neural buatan kabur. Set peraturan yang digunakan adalah kepuasan pelanggan dan uranium manakala set data yang digunakan ialah arak, iris, servo, kekuatan mampat konkrit dan uranium. Hasil untuk rangkaian dua fasa mendedahkan keunggulan kaedah ini berbanding dengan rangkaian konvensional satu fasa. Rangkaian dua fasa telah meningkat sebanyak 35% dan 98% untuk masa pelaksanaan dan 27% dan 26% untuk ralat umum. Hasil untuk rangkaian neural berasaskan GA didedahkan berbanding dengan rangkaian neural buatan jelas berasaskan GA. Sementara itu perbandingan dengan kaedah pengkomputeran granular yang lain yang berkaitan dijalankan menggunakan set data penanda aras iris. Hasil untuk rangkaian ini menunjukkan prestasi purata sebanyak 82.1%. Hasil daripada kaedah yang disarankan telah dianalisis dari segi statistik, kekuatan aturan dan pengkelasan tenaga menggunakan penanda aras set data perubatan. Oleh itu tesis ini menunjukkan bahawa rangkaian neural granular kabur berasaskan GA dan rangkaian neural buatan kabur berasaskan GA mampu sebagai taakulan bagi rangkaian neural buatan berdasarkan granul untuk kedua-dua data jelas dan tidak jelas.