4th Annual Research Symposium
HFMG Academic Affairs
(To be held on April 13, 2007 at Henry Ford Hospital, E&R Building)
Abstract Deadline: March 23, 2007

Data Mining from Epilepsy Database using PSO

Mostafa Ghannad-Rezaie[1], Hamid Soltanian-Zadeh[1,2], Kost V. Elisevich[3]

[1]Image Analysis Lab., Radiology Department, Henry Ford Health System, Detroit, Michigan, USA.
[2]Control and Intelligent Processing Center of Excellence, ECE Department, University of Tehran, Tehran, Iran.
[3]Department of Neurosurgery, Henry Ford Health System, Detroit, Michigan, USA.

Introduction: Improving healthcare quality, safety, and cost-effectiveness has stimulated interest in computer-based clinical decision support (CDS) systems. Despite recent achievements, successful systems are not easily replicated. The root problem is that although CDS is easy for on/off implementations, it is hard to make it robust. The apparent simplicity of an "*if...then…"* rule is seductive. Dynamic rule maintenance infrastructure is required to maintain and update rules. To make progress, it is essential that sufficient attention and resources are devoted to these processes and optimize them.

Methods: In the proposed method, a data driven metric is used to compare two cases in the database. The metric is defined base on critical case identification. By critical case, we mean the case with highest misunderstanding probability (known as support vector in the pattern recognition literature). Unlike traditional approaches, this metric could use knowledge of domain expert in form of "*if…then…*" rules or case similarity pairs. We use fast l_1 and l_p weighted heuristic optimization algorithms to find the most similar critical cases to a new case. Also, we maintain the rule set using a multi-agent optimization algorithm known as *particle swarm optimization* (PSO). The hybrid algorithm enables us to compensate small number of samples using domain knowledge in a systematic way. The simulation result shows that l_p optimization when $1<p<2$ establishes a reasonable trade-off between robustness and generalization.

Results and Conclusions: We applied the proposed approach to our epilepsy database called HBIDS. The database includes about 50 patients with tolerable missing values. We used similarity metric to establish a relation among patients and select surgery candidates. We compared our approach with well-know C4.5 and LOLIMOT data mining approaches. Subjective evaluation of the rule set shows higher robustness and more consistency with expert knowledge compared to the traditional algorithms. Moreover, in numerical evaluations, our algorithm outperforms its nearest competitor algorithm. The results are supportive of using critical cases to predict the outcome for a new case. The proposed method may be applicable to similar databases.