Journal of Software Engineering and Applications, 2011, 4, 388-390
doi:10.4236/jsea.2011.46044 Published Online June 2011 (http://www.SciRP.org/journal/jsea)

Prediction of Outcome in the Vegetative State by Machine Learning Algorithms: A Model for Clinicians?

Loris Pignolo[1*], Vincenzo Lagani[2]

[1]S. Anna Rehabilitation Institute and RAN-Research on Advanced Neuro-rehabilitation, Crotone, Italy; [2]Biomedical Laboratory, Institute of Computer Science, Foundation for Research and Technology-Hellas, Heraklion, Greece.
Email: l.pignolo@istitutosantanna.it, vlagani@forth.gr

Received June 8th, 2010; revised March 13th, 2011; accepted April 20th, 2011.

ABSTRACT

Purpose of this study was to compare different Machine Learning classifiers (*C*4.5, *Support Vector Machine*, *Naive Bayes*, *K-NN*) *in the early prediction of outcome of the subjects in vegetative state due to traumatic brain injury. Accuracy proved acceptable for all compared methods* (*AUC* > 0.8), *but sensitivity and specificity varied considerably and only some classifiers* (*in particular, Support Vector Machine*) *appear applicable models in the clinical routine. A combined use of classifiers is advisable.*

Keywords*: *Artificial Intelligence, Vegetative State, Clinical Outcome, Prognosis

1. Introduction

Vegetative state (VS) is a clinical condition resulting from severe brain damage and characterized by the lack of awareness (of self and environment), voluntary or otherwise purposeful behavioral responses, and communication in patients [1-6]. The early prediction of outcome has high priority for both the patient's family and attending physicians, who need to plan dedication and resources [7-9]. However, no prognostic model has yet proved suitable of generalization across different settings. Neither the traditional or newly developed electrophysiological techniques (waking and sleep EEG, visual auditory or somatosensory evoked potentials, event-related potentials, EEG brain imaging) nor other methods of functional evaluation (fMRI) are usable prognostic tools. An empirical prognostic model has been developed based on the appearance/disappearance and timing of observation of conventional neurological signs [10,11] and made available to clinicians. Eye tracking (*i.e.* smooth eye pursuit of moving target), spontaneous motility, the oculocephalic reflex and the chewing/sucking reflex were the relevant signs according to the model developed by data mining algorithms, with greater accuracy of prediction in post-traumatic subjects in VS. Prognosis nevertheless depended on evolution over time rather than on early assessment.

Purpose of this study is to assess the efficacy of theavailable decisional models in comparison with those suitable of application in this context. To this end, a subpopulation of subjects in VS due to head trauma was studied at admittance to the dedicated unit for VS of the Institute S. Anna – RAN.

2. Methods

Two hundreds forty one patients admitted to the dedicated semi-intensive care over an 11-year period (April 1998 - June 2009) were considered retrospectively. Inclusion criteria were posttraumatic aetiology and a diagnosis of vegetative state, as clinically defined compliant to the criteria suggested by the Multi-Society Task Force and the guidelines of the London Consensus Conference [5,12]. Patients recovering consciousness within 4 weeks from brain injury, with severe spinal fractures spinal cord damage or otherwise requiring intensive treatment or surgery, or in VS resulting from alcohol or drug overdose, or with a Glasgow Coma Scale [13] rating > 8 at admission were excluded in order to minimize misdiagnoses.

For each patient the following parameters were collected: age, sex, days in hospital reanimation unit, Gascow Coma Scale at admission, necessity of tracheotomy, necessity of nasogastric feeding tube (NFT), need of per-

cutaneous endoscopic gastronomy (PEG), presence of dysautonomic syndrome. The presence of preselected neurological signs (spontaneous motility, chewing and sucking reflexes, oculo-cephalic reflex, eye tracking) was searched for at admission. The outcome were assessed by the Glasgow Outcome Scale (GOS) [14]; patients were allocated to two main clusters of classes, *i.e.* to negative (GOS 1 or 2) or positive (GOS score 3 to 5) outcome.

All the numerical variables were normalized in the interval (0,1). Four different machine learning methods were used for data analysis: K Nearest Neighbours (K-NN) [15], Naïve Bayes [16], C4.5 Decision Tree [17,18], Support Vector Machine (SVM) [19]. The selected algorithms follow distinct approaches in model development. A major methodological issue was to compare the suitability of a decision tree rule-based approach (C4.5), Bayesian methods or numerical algorithms (SVM, K-NN) [20].

These algorithms require the "a priori" specification of one or more parameters [21]. We used a repeated cross validation procedure to evaluate the results of each parameter configuration and avoid over-fitting. Cross validation is an established technique for the assessment of the generalizability of the result of a statistical analysis [22]. It partitions the dataset in several complementary folds, to then use each fold as test-set, while the remaining dataset is used to build a model/execute the analysis. The results of the different test-sets are then averaged. The repetition of dataset splitting guarantees the results independence from the actual dataset subdivision [23].

As a further step, we used the Area Under the Curve (AUC) [24,25] metrics to identify the best configuration of parameters for each machine learning algorithm. AUC values can range from 0 (total misclassification) to 1 (perfect classification); random classification is arounf 0.5. AUC has the advantage of independency from distribution, *i.e.* evaluation of the results is not biase in over-represented classes. In addition, we estimated for each algorithm the sensitivity and specificity for positive outcome [26]. All the experiments were performed making use of the WEKA (Waikato Environment for Knowledge Analysis) software [27,28].

3. Results and Discussion

AUC values indicate good classification performances for all the used classifiers; however, the sensitivity and specificity estimates outline differences among the used statistical methods (**Tables 1** and **2**). Sensitivity was always higher than specificity, therefore indicating that the methods detect positive outcomes more easily (*i.e.* positive outcome can be excluded for patients classified as candidates to a negative outcome; on the contrary, low specificity does not exclude a negative prognosis for patients whose outcome was classified as positive); high sensitivity of the machine learning model is therefore a prerequisite in the early identification of patients with positive prognosis. In this respect, SVM seems to be the most suitable models for performing differential prognostic evaluations. On the contrary, Naïve Bayes sensitivity and specificity values are similar, indicating that such models can identify both the positive and negative outcomes with acceptable accuracy. C4.5 and K-NN methods proved Pareto dominated [29] by the other two algorithms.

Table 1. AUC for each dataset and algorithm.

	AUC
C4.5	0.84
SVM	0.81
Naive Bayes	0.91
K-NN	0.88

Table 2. Sensitivity and specificity for class "positive outcome" (GOS values 3, 4 and 5).

	sensitivity	specificity
C4.5	0.94	0.58
SVM	0.97	0.65
Naive Bayes	0.83	0.77
K-NN	0.97	0.44

4. Conclusions

We compared four different machine learning methods (C4.5, SVM, Naïve Bayes and K-NN) to identify the most suitable algorithm in the prognostic evaluation of subjects in vegetative state. All the tested algorithms are usable in this respect. SVM can be used for differential prognostic evaluations, *i.e.* SVM models may be a useful clinical tool to exclude a positive outcome. K-NN and C4.5 could be used for the same purpose, but their sensitivity and specificity are inferior to SVM. The Naïve Bayes classifiers do not appear usable for differential prognosis, due to poor efficiency in recognizing a specific class of subjects, but have limited classification errors and can be still considered as a valid (ancillary) prognostic tool. A decision to use classifiers specialized in differential prognosis or predictive models able to give reasonably accurate evaluation for both classes would depend on the clinical rationale and applied protocols. It may be worth noting that C4.5 remains a tool with potential clinical application in spite of poor performances; it is the only algorithm among those studied to be able to provide graphical models user-friendly to the clinician.

The ccombined use of different machine learning tools may be preferable in the vegetative state clinical setting,

with the approach complexity predictably compensated for by the overall resources that the care of patients in vegetative state requires.

REFERENCES

[1] B. Jennett and F. Plum, "Persistent Vegetative State after Brain Damage: A Syndrome in Search of a Name," *Lancet*, Vol. 1, 1972, pp. 734-619. doi:10.1016/S0140-6736(72)90242-5

[2] G. Dolce and L. Sazbon, "The Posttraumatic Vegetative state," Stuttgart, Thiene, 2002.

[3] S. Laureys, "The Neural Correlate of (un) Awareness: Lessons from the Vegetative State," *Trends in Cognitive Sciences*, Vol. 9, No. 12, 2005, pp. 556-559. doi:10.1016/j.tics.2005.10.010

[4] B. Jennett, "The Vegetative State," Cambridge University Press, Cambridge, 2002. doi:10.1017/CBO9780511545535

[5] Multi-Society Task Force on PVS, "Statement on Medical Aspects of the Persistent Vegetative State," *The New England Journal of Medicine*, Vol. 330, No. 21, 1994, pp. 1499-1508. doi:10.1056/NEJM199405263302107

[6] A. Zeman, "Consciousness," *Brain*, Vol. 124, 2001, pp. 1263-1289.

[7] R. Braakman, W. B. Jennett and J. M. Minderhoud, "Prognosis of the Post Traumatic Vegetative State," *Acta Neurologica Scandinavica*, Vol. 95, 1988, pp. 49-52. doi:10.1007/BF01793082

[8] E. Schmutzard, A. Kampfl, G. Franz, B. Pfausler, H. P. Haring, H. Ulmer, S. Felber, S. Golaszewski and F. Aichner, "Prediction of Recovery from Post Traumatic Vegetative State with Cerebral Magnetic-Resonance Imaging," *Lancet*, Vol. 351, No. 9118, 1998, pp. 1663-1671.

[9] A. Rovlias and S. Kotsou, "Classification and Regression Tree for Prediction of Outcome after Severe Head Injury Using Simple Clinical and Laboratory Variables," *Journal of Neurotrauma*, Vol. 21, 2004, pp. 886-893. doi:10.1089/0897715041526249

[10] G. Dolce, M. Quintieri, S. Serra, V. Lagani and L. Pignolo, "Clinical Signs and Early Prognosis: A Decisional Tree, Data Mining Study," *Brain Injury*, Vol. 22, No. 7-8, 2008, pp. 617-623. doi:10.1080/02699050802132503

[11] L. Pignolo, M. Quintieri and W. G. Sannita, "The Glasgow Outcome Scale in Vegetative State: A Possible Source of Bias," *Brain Injury*, Vol. 23, No. 1, 2009, pp. 1-2. doi:10.1080/02699050802595873

[12] K. Andrews (Chairman), "International Working Party Report on the Vegetative State," Royal Hospital for Neuro-Disability, 1996. http://www.comarecovery.org/s.htm

[13] G. Teasdale and B. Jennet, "Assessment of Coma and Impaired Consciousness: A Practical Scale," *Lancet*, Vol. 2, 1974, pp. 81-84.doi:10.1016/S0140-6736(74)91639-0

[14] B. Jennet and M. Bond, "Assessment Outcome after Severe Brain Damage: A Practical Scale," *Lancet*, Vol. 1, 1976, pp. 480-484.

[15] G. Shakhnarovish, T. Darrell and P. Indyk, "Nearest-Neighbor Methods in Learning and Vision," The MIT Press, Cambridge, 2005.

[16] P. Domingos and M. Pazzani, "On the Optimality of the Simple Bayesian Classifier under Zero-One Loss," *Machine Learning*, Vol. 29, 1997, pp. 103-137. doi:10.1023/A:1007413511361

[17] J. R. Quinlan, "C4.5: Programs for Machine Learning," Morgan Kaufmann Publishers, Waltham, 1993.

[18] L. Breiman, J. H. Friedman and R. A. Olshen, "Classification and Regression Trees," Wadsworth International, Belmont, CA, 1984.

[19] N. Cristianini and J. Shawe-Taylor, "An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods," Cambridge University Press, Cambridge, 2000.

[20] R. C. Holte, "Very Simple Classification Rules Perform Well on Most Commonly Used Datasets," *Machine Learning*, Vol. 11, No. 1, 1993, pp. 63-90. doi:10.1023/A:1022631118932

[21] F. Moser, G. E. Rong and E. Martin, "Joint Cluster Analysis of Attribute and Relationship Data without Apriori Specification of the Number of Clusters," *In Proceedings of the* 13*th ACM SIGKDD International Conference on Know- ledge Discovery and Data Mining*, San Jose, California, 12-15 August 2007.

[22] M. Stone, "Cross-Validatory Choice and Assessment of Statistical Predictions," *Journal of the Royal Statistical Society*, Vol. 36, No. 2, 1974, pp. 111-147.

[23] R. Picard and D. Cook, "Cross-Validation of Regression Models," *Journal of the American Statistical Association*, Vol. 79, No. 387, 1984, pp. 575-583. doi:10.2307/2288403

[24] M. L. Thompson and W. Zucchini, "On the Statistical Analysis of ROC Curves," *Statistics in Medicine*, Vol. 8, No. 10, 1989, pp. 1277-1290. doi:10.1002/sim.4780081011

[25] M. S, Pepe, "A Regression Modelling Framework for ROC Curves in Medical Diagnostic Testing," *Biometrika*, Vol. 84, 1998, pp. 595-608.

[26] R. W. Thatcher, R. A. Walker, C. J. Biver, D. M. North and R. Curtin, "Sensitivity and Specificity of an EEG Normative Database: Validation and Clinical Correlation," *Journal of Neurotherapy*, Vol. 7, No. 3-4, 2003, pp. 87-121. doi:10.1300/J184v07n03_05

[27] E. Frank, "Machine Learning with WEKA," University of Waikato, New Zealand, 1999. http://www.cs.columbia.edu/~jaxin/nlpmeetings/2004-04-22-lokesh.ppt http://ww w.cs.waikato.ac.nz/ml/weka

[28] J. H. Van Bemmel and M. A. Munsen, "Handbook of Medical Informatics," Springer-Verlag, Berlin, 1997.

[29] M. J. Osborne and A. Rubenstein, "A Course in Game Theory," MIT Press, Cambridge, 1994.