Intelligent Data Analysis in Medicine and Pharmacology

Intelligent Data Analysis in Medicine and Pharmacology

Edited by

Nada Lavrač
J. Stefan Institute, Ljubljana, Slovenia

Elpida Keravnou
University of Cyprus, Nicosia, Cyprus

Blaž Zupan
J. Stefan Institute, Ljubljana, Slovenia

Kluwer Academic Publishers
Boston/Dordrecht/London

Contents

Contributing Authors

Nada Lavrač is a research associate at the Department of Intelligent Systems, J. Stefan Institute, Ljubljana, Slovenia. She is coauthor of *KARDIO: A Study in Deep and Qualitative Knowledge for Expert Systems*, The MIT Press 1989, and *Inductive Logic Programming: Techniques and Applications*, Ellis Horwood 1994.

Dunja Mladenić is a final year PhD student in Computer Science at the Department of Intelligent Systems, J. Stefan Institute, Ljubljana, Slovenia. Her current research focuses on using machine learning in data analysis, with particular interest in learning from text applied on World Wide Web documents and intelligent agents.

Iztok A. Pilih is a trauma surgeon at the Department of Trauma Surgery, Celje General Hospital, Slovenia. He has completed his Master's Degree studying patients with severe head injury.

Tine S. Prevec is Professor of Neurology at the School of Medicine, University of Ljubljana, Slovenia. At the Department of Neurology in Innsbruck, Austria, he has studied the cognitive functions in patients recovering after a severe head injury.

1 DATA ANALYSIS OF PATIENTS WITH SEVERE HEAD INJURY

Iztok A. Pilih,
Dunja Mladenić,
Nada Lavrač,
Tine S. Prevec

Abstract: This chapter presents an application of decision tree induction to the problem of the prediction of outcome after a severe head injury. The study shows that induced decision trees are useful for the analysis of the importance of clinical parameters and of their combinations for the evaluation of the severity of brain injury and for outcome prediction. Due to a small number of patient data available for this study the induced decision trees cannot yet be considered as a reliable prognostic tool. Nevertheless, meaningful regularities have been discovered that help in the analysis of this difficult prognostic task.

1.1 INTRODUCTION

The prediction of the outcome of severe head injury is a difficult and complex problem since it depends on a large number of factors: those defining the condition of a patient before an injury, the ones acting during the injury, as well as the numerous factors acting in the period after the injury. Limitations of the early clinical evaluation of the brain injury (at the site of the accident or on admission to the hospital) and the long period of recovery make the problem even more demanding [Vollmer, 1993]. In spite of these limitations, the outcome prediction is needed and is useful for several reasons. It provides the insight into the pathophysiological events after head injury that makes therapeutic decisions easier, it is needed to inform the relatives, it helps in the evaluation of the effectiveness of the medical care and enables the timely selection of the most suitable rehabilitation program [Vollmer, 1993, Braakman, 1992].

Traditional approaches to the problem consist of a systematic collection of a large number of relevant factors, the statistical analysis of their prognostic values, and the search for the most valuable combinations of factors that influence the actual outcome after a severe head injury . Such approaches are based on large homogeneous populations of data of patients with severe head injury and on the application of statistical methods [Braakman, 1992]. Several prognostic models have been created by such approaches. They can provide more reliable prognosis than the lifelong personal experience [Braakman, 1992]. Prognostic models are usually presented as mathematical models, graphs or tables. Choi and co-workers [Choi, et al., 1991] presented a prognostic model in the form of a decision tree .

This work presents an application of machine learning [Quinlan, 1986, Michie, et al., 1994] to the problem of outcome prognosis after a severe head injury [Pilih, et al., 1995, Pilih, et al., 1996]. The chapter is organized as follows. Section 1.2 presents the prognostic problem: the available dataset, the prognostic classes and the clinical parameters are described. The machine learning methodology of decision tree induction is briefly described in Section 1.3. The experiments and results of decision tree induction are presented in the next three sections. Sections 1.4 and 1.5 deal with the problem of outcome prediction, whereas Section 1.6 deals with the analysis of the brainstem syndrome which is one of the most important parameters for outcome prediction. The chapter concludes with a summary and directions for further work.

1.2 PATIENT DATA

This study is based on the data of patients who were referred to the Celje General Hospital, Slovenia, after a severe head injury in the period from October 1992 till April 1993, first treated in the Intensive Care Unit and later at the

Department of Trauma Surgery. In this period, 38 patients were admitted to the hospital.

All the patients were submitted to a diagnostic protocol used as the doctrinary procedure for the patients with severe head injury at the Department of Trauma Surgery. In addition to the general clinical examination and standard laboratory tests, the diagnostic protocol includes computer axial tomography (CT), and the evaluation of coma according to the Glasgow coma scale (GCS) [Teasdal and Jennett, 1974]. The neurological evaluation of brainstem functions separating brainstem affection into seven brainstem syndromes (BSS) according to Gerstenbrand and co-workers [Gerstenbrand and Lücking, 1970, Gerstenbrand and Rumpl, 1983, Gerstenbrand, et al., 1990] was also applied.

Prediction classes: Glasgow outcome scale

The outcome was evaluated six months after the injury according to the Glasgow outcome scale (GOS). Patients were classified into five groups: GOS1 - fatal outcome, GOS2 - persistent vegetative state (patients with complete expression of apallic syndrome according to [Gerstenbrand and Rumpl, 1992] were included in this group), GOS3 - severely disabled patients needing much support for their personal care and daily activities due to their mental and physical handicaps, GOS4 - moderately disabled patients, able of taking care of themselves in their daily life, GOS5 - patients with a good outcome (able to continue their education, professional and social engagements similar as before the head injury) [Jennett and Bond, 1975].

Eighteen out of the 38 patients (47%) died in the first six months after the injury, their survival time being from 5 hours to 97 days. More than half of these patients (25% of all the patients) died during the first week. Six months after the injury, among the 20 patients that survived (53% of all the patients), one patient was in the persistent vegetative state (GOS2), three in the severe disability group (GOS3), five were moderately disabled (GOS4), and eleven had a good outcome (GOS5) .

Description of clinical parameters

Age. The mean value of the patients' age was 46 (standard deviation 19), within the range 16–81 years. In the group of the youngest patients ($\leq$ 20 years) most had a good outcome (66.7%), the death rate of patients aged under 40 was 18.7%, whereas above 40 it was much higher: 68.2%.

Computer axial tomography (CT) was performed by a Siemens Somatom HiQ. Only the CTs recorded upon admission were considered in this study. Six types of abnormalities were evaluated: epidural haematoma, subdural haematoma, intracerebral haematoma or contusion, subarachnoidal haemorrhage, intraventricular haemorrhage, and midline displacement larger than 5 mm. `CT score` was computed as a simple sum (the range is from 0 to 6) of these CT abnormalities. Five patients who had no CT abnormalities all survived with a good outcome, whereas the death rate of patients with more than one CT abnormalitiy was over 50%.

Glasgow coma scale (`GCS`). The score based on `GCS` is an estimate of the affection of consciousness. Verbal responsiveness (range 1–5), eye opening (range 1–4) and motor responses (range 1–6) are evaluated. The `GCS` score is computed as a sum of the above estimates, expressed in the range from 3 (the most severe affection) to 15 (no affection) [Teasdal and Jennett, 1974]. This study includes only patients with $\mathtt{GCS} \leq 8$ which we used as a definition of a severe head injury . The death rate of patients with `GCS` below 5 was 83.3%, whereas above 5 the death rate was much lower: 17.6%.

Brainstem syndromes (`BSS`) [Gerstenbrand and Lücking, 1970, Gerstenbrand and Rumpl, 1983, Gerstenbrand, et al., 1990] separate brainstem dysfunction into seven stages, from 1 to 7, where stage 7 corresponds to the slightest affection of the brainstem `MS1` (Mesencephalic Syndrome 1) , stage 6 is `MS2a`, stage 5 is `MS2b`, stage 4 is `MS3`, stage 3 is `MS4`, whereas stages 2 and 1 are Bulbar Syndromes 1 and 2 (`BS1` and `BS2`), respectively. `BS2` is the most severe brainstem affection; patients do not recover if this state lasts for more than 20 minutes [Gerstenbrand, et al., 1990]. The patients with `MS1` were not included in this study because they do not fulfil the criterion of a severe head injury ($\mathtt{GCS} \leq 8$) . Brainstem syndromes are estimated from the basic clinical signs (see below) [Gerstenbrand and Lücking, 1970, Gerstenbrand and Rumpl, 1983, Gerstenbrand, et al., 1990], in accordance with Table A1 given in the appendix.

Basic clinical signs. The basic clinical attributes used in the descriptions of patient records include [Gerstenbrand and Lücking, 1970, Gerstenbrand and Rumpl, 1983, Gerstenbrand, et al., 1990]: `Vigilance`; reaction parameters: `Reaction to acoustic stimuli`, `Reaction to pain stimuli` (also called reaction to nociceptive stimuli); optomotoricity parameters: `Position of bulbi`, `Motility of bulbi`, `Pupillary diameter`, `Pupillary reaction to light`; parameters of body motility: `Body posture`, `Muscle tone`, `Pyramidal signs`; vegetative parameters: `Respiration rate`, `Heart rate`, `Systolic blood pressure`, `Diastolic blood pressure`, `Body temperature`.

For illustration, we list the values of some clinical attributes: `Vigilance` (0 - coma, 1 - sopor, 2 - somnolence); `Reactions to pain stimuli` (0 - no motor reaction, 1 - rest of extension of upper and lower extremities, 2 - extension of upper and lower extremities, 3 - flexion of upper and extension of lower extremities, 4 - rest of nonfinalized defence reaction, 5 - delayed nonpurposeful defence reaction, 6 - prompt finalized defence reaction); `Body posture` (0 - flaccid posture, 1 - rest of extension of upper and lower limbs, 2 - extension of upper and lower limbs, 3 - flexion of upper and extension of lower limbs, 4 - mass movements of upper limbs and extended position of lower limbs, 4 - mass movements of upper limbs and extension trend in lower limbs, 5 - mass movements and rotations). Note that since two values of `Body posture` are very hard to be distinguished, they were merged into a single value (value 4). Such merging of values occured also in some other attributes (see Table A1 in the appendix).

1.3 INDUCTION OF DECISION TREES

In supervised machine learning , a set of examples with known classifications is given. An example is described by an outcome (class) and the values of a fixed collection of parameters (attributes). Each attribute can either have a finite set of values (discrete attribute) or take real numbers as values (continuous attribute). For instance, in our prediction problem the examples are patient records that comprise both continues (e.g., `Age`) and discrete attributes (e.g., `GCS`) .

Machine learning tools [Quinlan, 1986, Michie, et al., 1994] have been applied in a variety of medical domains to help to solve diagnostic and prognostic problems [Kononenko and Kukar, 1995]. These tools enable the induction of diagnostic and prognostic knowledge in the form of rules or decision trees .

In this study, the program Magnus Assistant [Mladenić, 1990] was used to construct decision trees . Magnus Assistant is a descendant of the program Assistant [Cestnik, et al., 1987] and belongs to the ID3 family of systems for top-down induction of decision trees [Quinlan, 1986]. The system recursively builds a binary decision tree. The nodes of the tree correspond to attributes, arcs correspond to values or sets/intervals of values of attributes, and leaves (terminal nodes) correspond to prognostic classes. In each recursive step of the decision tree construction, an attribute is selected and a subtree is built. Algorithm 1.3 gives an outline of the Magnus Assistant algorithm.

At each recursion step, the selection of the best attribute is based on the 'informativity' of the attribute [Quinlan, 1986] aimed at minimizing the expected number of tests needed for the classification of new cases. Tree construction is thus heuristically guided by choosing the 'most informative' attribute at the current node of a partially built tree. *Informativity* (usually referred to as *information gain*) measures the information gained by partitioning a set of

```
if all examples belong to the same class then
    this node becomes a leaf labelled with the class of examples
else
    if any of the pre-pruning criteria is satisfied then
        this node becomes a leaf labelled with the majority class of examples
    else
        for the next node select the best attribute (which minimizes the
            expected entropy of the training set)
        split the set of examples in the node into disjoint subsets according
            to binarized values of the attribute
        recursively repeat the whole algorithm for each subset
    end if
end if
```

Algorithm 1.1 Magnus Assistant algorithm.

examples according to the values of the attribute A:

$$info(A,E) = entropy(E) - \sum_v p_v(E) \times entropy(E_v)$$

where E denotes the set of examples on the current node of the decision tree, $p_v(E)$ denotes the probability of attribute value v in example set E, and E_v denotes the subset of examples for which attribute A has value v. The summation is over all the values of attribute A.

The definition of *entropy* is taken from the information theory and in our case it can be described as the average amount of information needed to identify the class of an example taken from the set of $Examples$:

$$entropy(E) = \sum_c p_c(E) \times \log_2 p_c(E)$$

In the formula, $p_c(E)$ stands for the probability of class c in a set of examples E and is usually calculated as the proportion of examples belonging to class c. The summation is over all the classes.

Magnus Assistant incorporates also mechanisms for dealing with noisy data (i.e., data comprising errors). The main idea underlying this mechanism is to 'prune' a decision tree [Cestnik, et al., 1987] in order to avoid overfitting the dataset of examples which may be erroneous. Magnus Assistant uses two kinds of tree pruning techniques to deal with noisy data : post-pruning and pre-pruning . Post-pruning is used after the tree induction to substitute some of the subtrees with terminal nodes (leaves). In post-pruning , the expected quality (predicted classification accuracy) of each internal node is compared with the expected quality of its subtrees to decide whether to prune the subtrees. On the other hand, pre-pruning is used during tree induction and enables noise handling by means of stopping the generation of a subtree and making a leaf node labelled with the majority class of the current examples. There are three pre-pruning criteria used by the above algorithm: (1) stop if the number of examples in the current node is smaller than the threshold value; (2) stop if the majority class in the current example set has a higher probability than the threshold probability (the so-called 'minimal weight threshold'); (3) stop if the informativity of the best attribute is below the threshold value. Threshold values for all the three pre-pruning parameters are set by the user and should reflect the expected level of noise in the data.

To classify a new case, a path from the root of the tree is selected on the basis of the values of attributes of the new patient to be classified. In this way, for a given patient record, the path leads to a leaf which assigns a class that labels the leaf. The selected path may be viewed as a generalization of the specific patient record for which the prediction is being determined. If a leaf is labelled with more than one class, each with the probability of class prediction, then the class with the highest probability is selected for the classification of a new case.

The entire decision tree reflects the detected regularities in the data, describing the properties that are characteristic for subsets of examples belonging to subtrees. The ordering of attributes (from the root towards the leaves of the tree) reflects also the importance of attributes for the outcome class in the leaf. The measure of attribute informativity is the selected measure of importance.

1.4 OUTCOME PREDICTION FOR TWO OUTCOME CLASSES

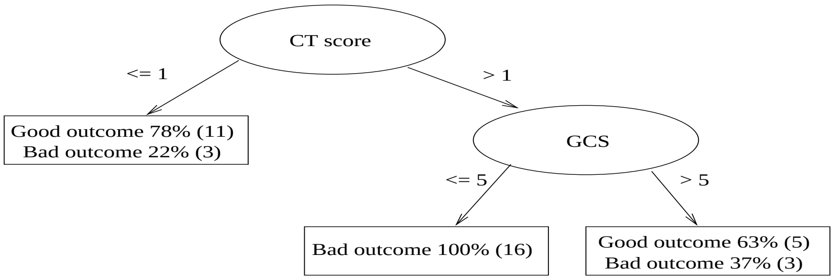
The main goal of these experiments was to predict the outcome six months after a severe head injury based on clinical findings on admission, and to study the prognostic values of the attributes. For this purpose, the Magnus Assistant program was used. Experiments were done by employing different tree pruning parameters. The two decision trees reported in this section were generated by Magnus Assistant using pre-pruning (minimal weight threshold = 40%).

Due to the small number of training examples available (38), we first decided to reduce the number of attributes to three only (either `GCS` or `BSS`, `CT score`, and `Age`) and the number of outcome categories to two: `GOS1`, `GOS2` and `GOS3` were grouped into the `Bad outcome` category, and `GOS4` and `GOS5` into the `Good outcome` category.

Experiments using GCS, CT score and Age

In this experiment, `GCS`, `CT score` and `Age` were the only attributes used. The decision tree induced by Magnus Assistant is shown in Figure 1.1.

Figure 1.1 Decision tree generated with three attributes (`GCS`, `CT score`, `Age`) and two outcome classes (`Bad`, `Good`). The percentages indicate the probabilities of class assignment, and the numbers in parentheses the number of patients in each outcome class.



`CT score` was selected by Magnus Assistant as the most informative attribute among the three attributes used. It was therefore used in the root of the tree, and the patients were split into two subsets, according to `CT score` values. Eleven out of 14 patients with `CT score` 0 or 1 had a good outcome. Patients with `CT score` > 1 were separated according to `GCS`. All the 16 patients with `GCS` ≤ 5 had a bad outcome. Patients with a higher `GCS` had better, but still uncertain outcomes.

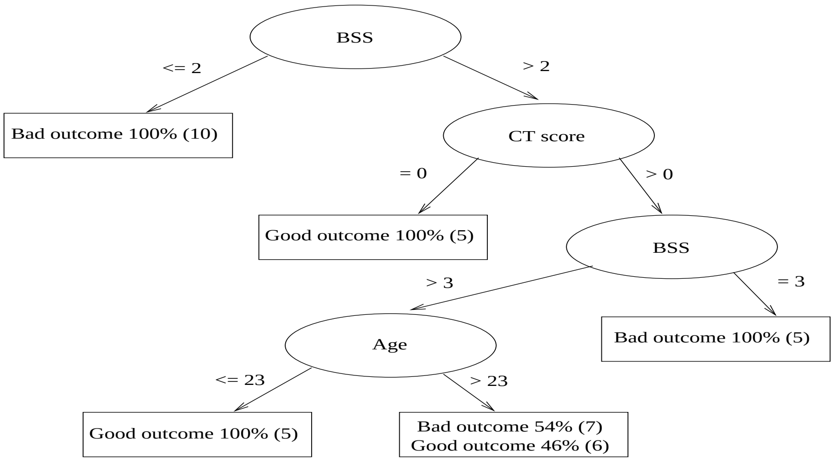
In this decision tree only 16 patients out of 38 are located in the leaf with only one outcome category and 22 patients belong to the leaves with both outcome categories. The impurity of the leaves is unfavourable. When a tree is used for classification, such leaves provide the classification as if the leaves were labelled with the most probable class, according to the probability of this class in the leaf.

In order to test the prognostic accuracy of the induced decision tree, we used the so-called *leave-one-out method*: the outcomes predicted by decision trees generated from the data of 37 patients were compared with the actual outcome of the patient that was eliminated from the training set of examples. Thirty-eight such tests were made. The proportion of correct classifications represents the measure of prognostic accuracy . Decision trees using the selected three parameters and two outcome classes had a relatively high prognostic accuracy (82%).

Experiments using BSS, CT score and Age

In these experiments, `BSS` was used instead of `GCS`. The decision tree generated with three attributes (`BSS`, `CT score` and `Age`) and two outcome classes (`Bad` and `Good`) is shown in Figure 1.2.

Figure 1.2 Decision tree generated with three attributes (`BSS`, `CT score`, `Age`) and two outcome classes (`Bad`, `Good`).



The splitting of the decision tree root is based on `BSS` which is the most informative attribute among the three attributes used. All the 10 patients with the most severe brainstem dysfunctions (values ≤ 2: 1 - `BS2` and 2 - `BS1`) died. The other 28 patients were further separated according to `CT score`. Five patients with `CT score` 0 had good outcomes, whereas patients with CT abnormalities had worse outcomes, and were further separated according to `BSS`. All five patients with the most severe midbrain syndrome (value 3 - `MS4`) died. Patients with a less severe brainstem disfunction (values > 3: `MS3`, `MS2b`, `MS2a`) were further separated according to `Age`: five patients with `Age` ≤ 23 had good outcomes. The outcome of older patients was worse and more uncertain.

In this decision tree, 25 patients (out of 38) are located in the leaves with only one outcome category and 13 patients belong to the leaf with both outcome categories.

The decision trees induced in this experiment again have a relatively high prognostic accuracy (79%) when tested by the leave-one-out method.

Analysis of results

These experiments show that induced decision trees are useful for the analysis of the importance of clinical parameters and their combinations for the prediction of the outcome. BSS has the highest informativity . It is followed by the sum of CT abnormalities. Next is the GCS score, and the last is the age of patients. Similar observations were discussed in other studies [Benzer, et al., 1987, Marosi, et al., 1991, Rumpl, et al., 1991]. Moreover, nonparametric correlation tests applied on the same dataset of 38 patients [Pilih, et al., 1995] also show that BSS has the highest correlation value, followed by GCS, CT score and the age of patients.

In the decision trees, where outcome prediction is based on BSS, 25 out of 38 patients (66%) are in the leaves with only one outcome class (good or bad). On the other hand, in the decision trees where outcome prediction is based on GCS there are only 16 out of 38 patients (42%) in the leaves with a single class outcome. We believe that an important factor influencing the impurity of leaves of GCS-based decision trees is a consequence of the poor discrimination of GCS. Other authors [Benzer, et al., 1987, Marosi, et al., 1991, Rumpl, et al., 1991] share this opinion. Marosi and co-authors [Marosi, et al., 1991] suggested to supplement the GCS data with the information about the pupillary diameter and reaction to light, as well as the position and motility of bulbi. These (and some other) data are already included in the classification of BSS, which (in our opinion) leads to the advantage of BSS over GCS.

The prognostic accuracy of decision trees based on BSS and GCS is similar: 79% and 82%, respectively. This result is similar to the result of Choi et al. [Choi, et al., 1991] who analyzed 555 patients with severe head injury with similar methods and have reported about 77.7% of correct prognoses. However, the similar amount of correct predictions based on BSS and GCS is somewhat unexpected since there is a larger amount of 'pure' predictions in the leaves of decision trees based on BSS. This is probably mostly due to the limited number of outcome classes (two classes only) used for this analysis. In part, this may also be due to the small number of patients in our study, where a single patient may significantly influence the structure of the decision tree.

1.5 OUTCOME PREDICTION FOR FIVE OUTCOME CLASSES

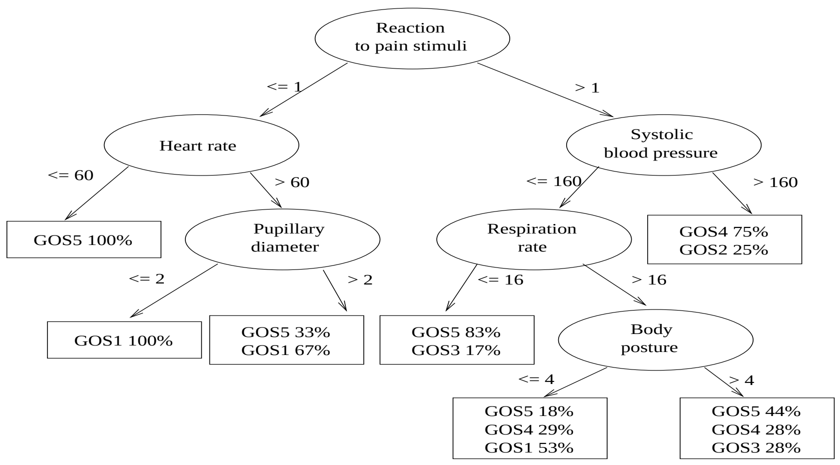
These experiments were mainly designed to test whether it is possible to replace the complex BSS parameter with the basic clinical parameters that are used for BSS estimation.

Experiments using basic clinical signs

In the first experiment only the basic clinical attributes were used (`BSS` was not used). The decision tree, generated by Magnus Assistant using pre-pruning (minimal weight threshold = 20%), is shown in Figure 1.3.

The attribute `Reaction to pain stimuli` is used in the root of the decision tree since it has the highest informativity . According to this attribute, patients are divided into two subgroups: the patients with poor or no motor reaction (values 1 and 0) and the others with a better reaction (values 2 or more). Patients in the first left branch were further separated according to `Heart rate`. Those with a lower heart rate (60 beats/minute or less) had a good outcome (one patient), whereas patients with `Heart rate` > 60 mostly died: out of 16 patients 15 died (`GOS1`) and one survived (`GOS5`) . The patients with better motor reactions to pain (values 2 or more) in the first right branch of the tree (21 patients) were further separated on the basis of the attribute `Systolic blood pressure`. The attribute `Respiration rate` was used in the next node of the tree. Patients with the respiration rate up to normal (16 respirations/minute or less) all survived (most with a good outcome and one severely disabled). The patients with higher respiration rates were separated in the next node according to the position of the body. The patients who showed mass movements and rotations had better outcomes than the others.

Figure 1.3 Decision tree generated with basic clinical attributes and five outcome categories.



Experiments using basic clinical signs and BSS

In the second experiment, `BSS` was used in addition to the basic clinical attributes. Unexpectedly, the generated decision tree was identical to the one generated without using `BSS` . Therefore, Figure 1.3 represents also the decision tree generated in this experiment.

When testing the informativity of the basic clinical attributes and `BSS` , `BSS` and `Reaction to pain stimuli` turned out to have an equal information score. This is the reason why the structure of the decision tree does not change if the root node `Reaction to pain stimuli` is replaced by `BSS` . However in this case, the left and right arcs leading from the root `BSS` are labelled by values ≤ 2 and > 2, respectively.

1.6 BRAINSTEM SYNDROME DETERMINATION

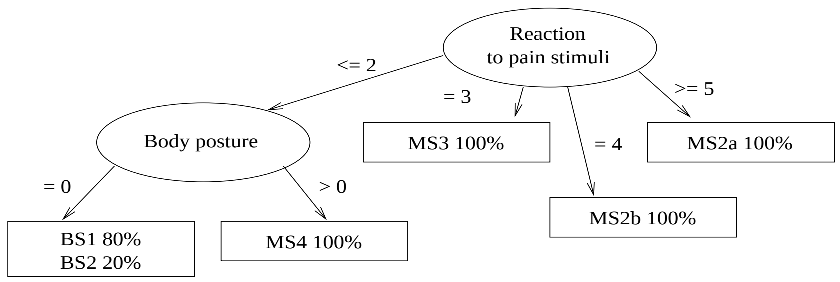
As shown in Sections 1.4 and 1.5, brainstem syndromes (`BSS`), as well as `Reaction to pain stimuli` are the most important for outcome prediction. However, it is sometimes difficult to determine `BSS` because of the numerous clinical signs that define them. In our study it also occurred that different clinical signs had values that belong to different brainstem syndromes . In such a situation it is not straightforward how to determine the actual `BSS`. Moreover, the searching for all the clinical signs which are needed for a formally complete determination of `BSS` (see Table A1 in the appendix) is also time consuming and is problematic in urgent situations at the admission of heavily injured patients when time, measured in minutes, is one of the most decisive factors for the outcome. Consequently, the experiments described in this section were intended to modify `BSS` determination by possibly reducing the present complex set of signs listed in Table A1 to the most essential (informative) ones for `BSS` determination, when the available time is limited.

To achieve this goal, two experiments for determining `BSS` were designed: the first one using the actual patient data for decision tree construction, and the second by generating training examples from the data about the main neurological signs of Table A1. The two decision trees for determining `BSS` were induced by Magnus Assistant using pre-pruning (minimal weight threshold = 20%).

Using the actual data of patients

This experiment was based on the available data of 38 patients and the `BSS` determination of the trauma surgeon, specially trained in head trauma. The induced decision tree for `BSS` determination is shown in Figure 1.4. Note that, for the reason of compactness, the tree has been rewritten by merging some of the branches.

Figure 1.4 Decision tree for BSS determination using the actual patient data.



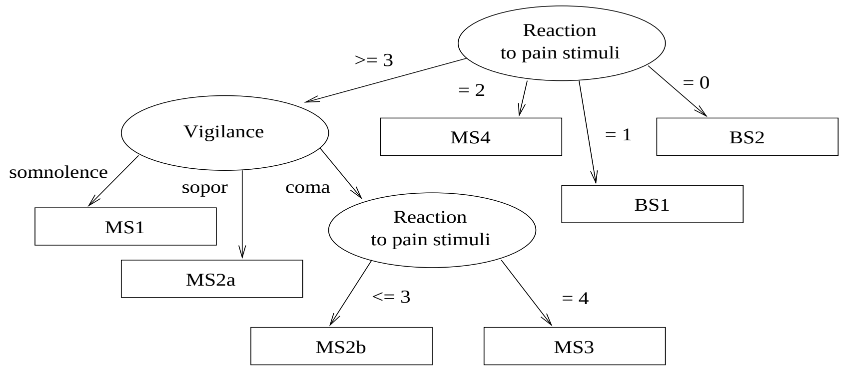
The analysis suggests that `Reaction to pain stimuli` and `Body posture` have the highest informative values for `BSS` determination. So both of them are used in the induced decision tree . The attribute `Reaction to pain stimuli` represents the root of the tree. It separates patients into midbrain syndromes `MS2a`, `MS2b` and `MS3`. The other patients are separated according to `Body posture` into `MS4` and into a common category comprising both bulbar syndromes `BS1` and `BS2`.

Using the basic clinical signs

In this experiment, Table A1 served as a source of training examples for `BSS` determination from the data about the main neurological signs determining `BSS` [Gerstenbrand, et al., 1990]. Seven training examples were generated from Table A1, one for each brainstem syndrome (an example is a table column). The induced decision tree is shown in Figure 1.5. Again, the tree has been rewritten by merging some of the branches in order to get a simpler tree representation.

The attribute `Reaction to pain stimuli` again turns out to be the most important parameter for determining the brainstem syndromes . The induced decision tree determines both bulbar syndromes (`BS1` and `BS2`) and the midbrain syndrome 4 (`MS4`) solely on the basis of `Reaction to pain stimuli`. Less severely affected patients are further separated on the basis of `Vigilance`. Patients that show somnolence on admission are classified as `MS1`, and those with sopor as `MS2a`. For the patients with coma, `Reaction to pain stimuli` was considered again in order to distinguish between `MS2b` and `MS3`.

Figure 1.5 Decision tree for BSS determination from the main neurological signs.



Analysis of results

The comparison of BSS determination using the two trees in Figures 1.4 and 1.5 shows some differences. The same syndrome was determined in 33 out of 38 patients (87%) using the actual patient data for tree construction, and in 23 out of 38 patients (61%) when the main neurological signs were used for tree construction. This speaks in favour of the decision tree in Figure 1.4 despite the fact that this decision tree (constructed from the actual patient data) joins BS1 and BS2 into a joint bulbar syndrome category (BS1 and BS2).

Most differences in decisions made by the two decision trees occur in the determination of the bulbar syndromes . In eight cases where the trauma surgeon determined BS1, the decision tree in Figure 1.5 proposes a more severe affection BS2. For three patients with MS2a evaluated by the trauma surgeon, the decision tree in Figure 1.5 proposes MS2b. For two other patients both trees determine BS2 instead of MS4. The analysis shows that in all the instances where the differences in classifications occur, the brainstem syndromes determined by decision trees are more severe than the ones determined conventionally using Table A1, despite the fact that also the conventional BSS assignment is based on the same knowledge source.

1.7 SUMMARY

This chapter presents an application of decision tree induction to the problem of the outcome prediction six months after a severe head injury . The study shows that induced decision trees are useful for the analysis of the importance of clinical parameters and of their combinations for the prediction of the outcome.

Among the parameters studied, the brainstem syndromes (BSS) turn out to be the most informative . However, the experiment in which BSS was replaced by the basic clinical attributes from which BSS is estimated has shown that BSS can successfully be replaced by `Reaction to pain stimuli` in the specific conditions of this experiment.

The use of decision tree induction methods turns out to be useful for improving the insight into the complex problem of severe head injury . The nature of the traumatic brain disease [Grčević, 1988], where different factors influence the course of recovery, makes the problem very difficult. By using the attributes available upon admission of patients included in this study, the precision level of about 80% can be achieved, when considering two outcome classes. Results of previous studies, based on often much larger groups of patients than ours, have also almost all reached a similar prediction accuracy. It seems that this is the approximate prediction accuracy of the data available on admission of patients with severe head injury . The rest of the outcome variability depends on the influence of factors guiding the course of the traumatic brain disease [Grčević, 1988]. Only the consideration of these factors by following the course of recovery of patients after severe head injury can complete the information needed to increase the reliability of the outcome prediction.

In further work, new cases will be added to the patient record database. The experiments in outcome prediction using the enlarged dataset will be repeated and the results compared. We are also planning to provide new training examples for the BSS estimation experiments in order to derive simple rules for fast BSS determination, that can, despite their simplicity, provide for more reliable diagnosis in time limiting conditions on admission to a hospital.

Acknowledgments

This work was funded by the Ministry of Science and Technology of Slovenia.

References

[Benzer, et al., 1987] Benzer, A., Mitterschiffthaler, G., Prugger, M., and Rumpl, E. (1987). Innsbruck Coma Scale versus Glasgow Coma Scale. *Notfallmedizin*, 13: 41–50.

[Braakman, 1992] Braakman, R. (1992). Early prediction of outcome in severe head injury. *Acta Neurochir.*, 116: 161–163

[Cestnik, et al., 1987] Cestnik, B., Kononenko, I., and Bratko, I. (1997). ASSISTANT 86: A knowledge elicitation tool for sophisticated users. In Bratko, I., Lavrač, N., editors, *Progress in Machine Learning*, pages 31–45. Sigma Press, Wilmslow.

[Choi, et al., 1991] Choi, S.C., Muizelaar, J.P., Barnes, T.Y., et al. (1991). Prediction tree for severely head-injuried patients. *J. Neurosurg.*, 75: 251–255.

[Gerstenbrand and Lücking, 1970] Gerstenbrand, F. and Lücking, C. (1970). Die akuten traumatischen Hirnstammschaden. *Arch. Psychiat. Nervenkrankh.*, 213: 264–281.

[Gerstenbrand and Rumpl, 1983] Gerstenbrand, F. and Rumpl, E. (1983). Das prolongierte Mittelhirnsyndrom traumatischer Genese. In Neumärker, K.J., editor, *Hirnstammläsionen. Neurologische, psychopatologische, morphologische, neurophysiologische and computertomographische Aspekte*, pages: 236–248. Hirzel, Leipzig.

[Gerstenbrand, et al., 1990] Gerstenbrand, F., Saltuari, L., Kofler, M., and Formisano, R. (1990) Clinical evaluation of severe head injury. *Neurologija*, 39: Suppl. 1: 71–88.

[Gerstenbrand and Rumpl, 1992] Gerstenbrand, F. and Rumpl, E. (1992). Apallisches Syndrom. In Hornbostel, H., Kaufmann, W. and Siegenthaler, editors, (1992). *Innere Medizin in Praxis und Klinik*, pages 7.63–8. Georg Thieme Verlag, Stuttgart, New York.

[Grčević, 1988] Grčević, N. (1988). Head injury. The concept of inner cerebral trauma. *Scand. J. Rehab. Med.*, Suppl. 17: 25–31.

[Jennett and Bond, 1975] Jennett, B. and Bond, M. (1975). Assessment of outcome after severe brain damage. *Lancet*, 1: 480–487.

[Kononenko and Kukar, 1995] Kononenko, I. and Kukar, M. (1995) Machine learning for medical diagnosis. In Lavrač, N., editor, *Computer Aided Data Analysis in Medicine*, pages 9–30. IJS Scientific Publishing, Ljubljana.

[Marosi, et al., 1991] Marosi, M.J., Gerstenbrand, F., Luef, G.J., Zimmermann, S., and Schmutzhard, E. (1991). Innsbruck coma rating scale. *Lancet*, 337: 1043.

[Michie, et al., 1994] Michie, D., Spiegelhalter, D.J., and Taylor, C.C., editors, (1994). *Machine Learning, Neural and Statistical Classification.* Ellis Horwood, Chichester.

[Mladenić, 1990] Mladenić, D. (1990). *The learning system Magnus Assistant.* BSc Thesis. Faculty of Computer and Information Sciences, University of Ljubljana (in Slovenian).

[Pilih, et al., 1995] Pilih, I.A., Mladenić, D., Lavrač, N., and Prevec, T.S. Using decision trees to predict the outcome after a severe head injury. In Lavrač, N., editor, *Computer Aided Data Analysis in Medicine*, pages 221–233, IJS Scientific Publishing, Ljubljana (in Slovene).

[Pilih, et al., 1996] Pilih, I.A., Mladenić, D., Lavrač, N., and Prevec, T.S. (1996) Data analysis of patients with severe head injprediction with decision trees. In Lavrač, N., Keravnou, E. and Zupan, B., editors, ECAI-96 Workshop Proceedings *Intelligent Data Analysis in Medicine and Pharmacology*, pages 69–72.

[Quinlan, 1986] Quinlan, J.R. (1986) Induction of decision trees. *Machine Learning*, 1: 81–106.

[Rumpl, et al., 1991] Rumpl, E., Gulle, H.D., Stadler, C., Millonig, H., and Prugger, M. (1991). Die Innsbrucker Koma-Skala beim schweren Schädel-Hirn-Trauma. *Intenziv- und Notfallbehandlung*, 16: 7–13.

[Teasdal and Jennett, 1974] Teasdale, G. and Jennett, B. (1974). Assessment of coma and impaired consciousness: A practical scale. *Lancet*, 2: 81–84.

[Vollmer, 1993] Vollmer, D.G. (1993) Prognosis and outcome of severe head injury. In Cooper, P.R., editor, *Head injury*, 3rd edition, pages 553–581, Williams & Wilkins, Baltimore.

Appendix: Clinical characteristics of brainstem syndromes

Table A1 presents seven brainstem syndromes described in terms of basic clinical signs. This table, taken from [Gerstenbrand, et al., 1990], is slightly modified and rewritten for the needs of our study: for instance, attribute values that are hard to be distinguished are merged into a single value, which increases the robustness and simplicity of induced trees.

Table A1 Clinical characteristics of brainstem syndromes

	7 - MS1	6 - MS2a	5 - MS2b	4 - MS3	3 - MS4	2 - BS1	1 - BS2
Vigilance	2 somno.	1 sopor	0 coma	0 coma	0 coma	0 coma	0 coma
Reaction to acoust. stimuli	1 slightly delayed	0	0	0	0	0	0
Reaction to pain stimuli	6 prompt finaliz. defence reaction	5 delayed non purpose. defence reaction	4 rest of a non finaliz. defence reaction	3 flexion of upper extens. of lower limbs	2 ext. of upper and lower limbs	1 rest of extens. of upper and lower limbs	0 no
Position of bulbi	2 normal	1 slight. diverg.	0 diverg.	0 fixed diverg.	0 fixed diverg.	0 fixed diverg.	0 fixed diverg.
Motility of bulbi	3 swaying	2 oscil.	1 disconj.	0 no	0 no	0 no	0 no
Pupillary diameter	3 normal	3 normal	3 normal	2 small	2 small	1 very large	0 large
Pup. reac. to light	2 normal	2 normal	2 normal	1 weak	1 weak	1 weak	0 no
Body posture	5 mass movem. and rotat.	4 mass movem. of upper, extens. trend in lower limbs	4 mass movem. of upper, extended position of lower limbs	3 flexion of upper, extens. of lower limbs	2 extens. of upper and lower limbs	1 rest of extens. of upper and lower limbs	0 flaccid posture
Muscle tone	4 normal	3 slight. incr. in lower limbs	3 incr. in lower limbs	1 incr.	2 mark. incr.	1 slight. incr.	0 flaccid
Respir. rate	4 normal	4 normal	3 Cheyne Stokes	2 rapid regular hyper-ventil.	2 regular hyper-ventil.	1 ataxic	0 no
Heart rate	1 slight. incr	3 normal	1 incr.	1 incr.	2 mark. incr.	1 incr.	0 decr.
Blood pressure	1 normal	1 normal	1 normal	2 slight. incr.	2 incr.	1 normal	0 decr.
Body temp.	3 normal	3 normal	1 slight. incr.	1 incr.	2 mark. incr.	1 incr.	0 decr.