Pre-publication draft of a paper which appeared in the Proceedings of the First Japanese Knowledge Acquisition for Knowledge-Based Systems Workshop (JKAW '90), Ohmsha, Tokyo

# Rule optimisation and theory optimisation:
# Heuristic search strategies for data driven machine learning

Geoffrey I. Webb

*Department of Computing and Mathematics,*
*Deakin University, Geelong, Australia*

## Abstract

Previous implementations of the Aq algorithm have used rule optimisation search strategies to attempt to develop optimal classification procedures. These strategies involve generating successive characteristic descriptions each of which is individually of maximal value. This is contrasted with theory optimisation search strategies which, instead, generate successive complete classification procedures from which those with the maximal value are selected. These two strategies have been applied to the domain of the diagnosis of Immunoglobulin A Nephropathy disease. The theory optimisation strategy was observed to out perform the rule optimisation strategy.

## 1 Introduction

The two key components of a data driven machine learning system are its *classification language* and its *search algorithm*. The classification language is a language in which classification procedures can be expressed. This defines the *classification procedures* that the system can produce. Typical examples of classification languages are predicate logic [1] and decision trees [2].

The search algorithm takes as input a *training set*, a set of examples of correct classifications of instances from a domain. The training set is used to select one of the classification procedures that can be described by the classification language. The algorithm attempts to select a classification procedure that will correctly classify all instances from the domain from which the training set is drawn.

A search algorithm usually attempts to maximize two measures -

- the number of instances from the training set that are correctly classified by the classification procedure; and
- the intrinsic value of the classification procedure.

Intrinsic value is primarily based on the simplicity of the classification procedure. It is generally accepted that the more simple classification procedure is preferable to the more complex [3].

The relative weights of these two measures reflect the degree of noise that is expected in the data. If no noise is expected then intrinsic value is usually used as a secondary measure to select between theories that classify the example set equally well.

Measures of the complexity of a classification procedure are not self-evident. For example, is a predicate logic statement in conjunctive normal form containing a two disjuncts that each contain three conjuncts more complex than a statement containing three disjuncts each of which contains two conjuncts? What are the relative complexities of two predicates (such as male(X) and olderthan(X, 25))?

Criteria for assigning the intrinsic value of a classification procedure and for weighing the intrinsic and the accuracy with relation to the training set are called *classification procedure evaluation criteria.*

It might be thought that, given a classification procedure evaluation criteria, the rest is trivial- a system need only search the space of classification procedures to find the procedure that maximizes the criteria. This is indeed possible for attribute-value machine learning where all attributes are discrete [4]. It is also possible where no disjunction is required [5]. However, where disjunction is required and ordinal attributes or structural relations are allowed by the language, the number of potential theories usually precludes exhaustive search. As a result, heuristic search strategies are usually employed.

The Aq algorithm [6] is a very successful data driven machine learning algorithm. It can be described as follows. The input is a set of instances divided into POS (instances belonging to the class of interest) and NEG (instances not belonging to the class of interest.) The output is a set of characteristic descriptions for the class.

```
Algorithm: Aq
Input:  POS (a training set of instances belonging to the class of interest)
        NEG (a training set of instances not belonging to the class of interest)
        criteria (classification procedure evaluation criteria)
Output:         R (a set of characteristic descriptions for the class.)
initialize R to empty.
while POS is not empty
        randomly select an instance i from POS
        select a characteristic description c that covers i and maximizes criteria.
        remove from POS all instances covered by c.
        add c to R
end while.
simplify R using both general and domain specific rules.
```

To obtain a complete classification procedure for a domain, the above algorithm is applied to each class in the domain, with *POS* and *NEG* set accordingly. Any new instance covered by a characteristic description so formed is deemed to belong to the class for which that characteristic description was developed.

It should be noted that the above is not a strict expression of the Aq algorithm which specifies that all most general characteristic descriptions that cover *c* and cover no instances in *NEG* should be generated at step 2, and that step 6 should select the subset of the characteristic descriptions so generated that maximizes the evaluation criteria. However, as noted above, an exhaustive search, which is required by this latter approach, is not computationally feasible for many domains. As a result, the exhaustive search is replaced in practice by the above algorithm with a heuristic search employed at step 2 that attempts to find the same characteristic description as would be found by an exhaustive search.

The heuristic search usually employed is a *beam search*. A beam search maintains a set of a maximum size, *n*, which contains the best *n* characteristic descriptions discovered to date. This set is called the beam. Successive cycles of the search attempt to refine the candidates in the beam, replacing current candidates with new candidates as better characteristic descriptions are discovered.

Two general strategies are usually employed. The beam may be seeded by the set of most general characteristic descriptions that cover the selected example. The search then considers successive specializations of candidates in the beam [7]. Alternatively, the beam may be seeded with the most specialized characteristic description that covers the selected example. In this ease the search considers successive generalizations of candidates in the beam [8].

A version of the latter strategy is the ROA algorithm (below.) Note that this algorithm constrains the search by only considering generalizations that expand the number of instances covered. This is achieved by only generalizing a rule by a *least generalization* [9] that causes the characteristic description to cover a new positive instance.

It can be seen that Aq attempts to construct the best theory by attempting to select successive characteristic descriptions each of which individually maximizes the given evaluation criteria. This is a *rule optimisation* search strategy. This type of strategy is subject to the flaw that individual characteristic descriptions that rate highly will not necessarily form a classification procedure that rates highly.

An alternative strategy is to generate successive complete classification procedures and to select that which maximizes the evaluation criteria. This is a *theory optimisation* search strategy. The TOA algorithm (below) is a modification of the ROA algorithm that utilizes theory optimisation.

It should be noted that the strict Aq algorithm, which generates all most general characteristic descriptions that cover each case examined, is a theory optimisation strategy, as an optimal theory is selected from the set of characteristic descriptions that are generated. The strict Aq algorithm differs from the TOA algorithm, however, in. that the TOA algorithm generates successive complete theories and selects between these. This strategy proves to be computationally efficient even with respect to continuous attributes.

Note that when *width* = 0, both the ROA and TOA algorithms perform identically.

```
Algorithm: ROA (Rule optimisation algorithm)
Input:  POS (a training set of instances belonging to the class of interest)
        NEG (a training set of instances not belonging to the class of interest)
        criteria (an evaluation criteria)
        width (an integer specifying the maximum width of the beam)
Output: R (a set of characteristic descriptions for the class.)
initialise R to empty.
while POS is not empty
        randomly select an instance i from POS.
        initialise beam with the most specific characteristic description that covers i
        for x = each instance in POS other than i, in random succession
           initialise beam* to an empty list.
           for y = each characteristic description in beam in succession
                 add the least generalization of y that covers x to beam*.
           end for.
           add beam* to beam.
        if beam contains more than width characteristic descriptions remove those that
        rate lowest according to criteria until only width characteristic descriptions remain.
        end for.
        remove from POS all instances covered by c.
        add c to R
end while.
simplify R using both general and domain specific rules.
```

# 2 Evaluation

To compare the algorithms they were implemented on a Sun 4/60 SPARCstation 1 and applied to the diagnosis of Immunoglobulin A Nephropathy disease. This domain was selected because it poses a difficult data driven machine learning task. Each instance is described by 37 attributes specifying relevant clinical information. For only 17 of these attributes are the values known for all instances. Only 276 instances are available. For no instance are all attribute values known. Of the 276 instances, 57 are positive and 219 are negative.

The data includes all patients for whom biopsies were conducted for renal conditions at Geelong Hospital from 1979 to 1989 inclusive. Clinical evaluation of a biopsy results in a definitive and objective diagnosis. Due to the objective nature of both the diagnosis and the attributes, the absence of noise from the data can be guaranteed.

```
Algorithm: TOA (Theory optimisation algorithm)
Input:   POS (a training set of instances belonging to the class of interest)
         NEG (a training set of instances not belonging to the class of interest)
         criteria (an evaluation criteria)
         width (an integer specifying the maximum number of theories to examine)
Output: r (a set of characteristic descriptions for the class)
initialise r to empty.
repeat width times
  initialise r* to empty.
  while POS is not empty
    randomly select an instance i from POS
    initialise c to the most specific characteristic description that covers i.
    for x = each instance in POS other than i, in random succession
         set c* to the least generalization of c that covers x.
         if the value criteria assigns to c* is greater than the value criteria assigns to c set c
         to c*
    end for.
    remove from POS all instances covered by c.
    add c to r
  end while.
  restore POS to its initial value.
  if r* rates higher than r according to criteria set r to r*.
end repeat.
simplify r using both general and domain specific rules.
```

Note that the algorithms as presented do not specify the evaluation criteria that are to be employed within a search. In order to test the algorithms, they were executed using simple evaluation criteria. The intended objective of these criteria is to develop the simplest possible classification procedure that correctly classifies all instances in the training set. Simplicity is measured by the number of characteristic descriptions in the classification procedure.

The algorithm used to implement these criteria is listed below. Essentially, of any two classification procedures, one with no negative cover is preferred to one with negative cover. If neither has negative cover then the one with the highest positive cover is preferred. Of any two classification procedures with no negative cover and identical positive cover, the one with the fewest characteristic descriptions is preferred. Note, $|C|$ denotes the number of characteristic descriptions in $C$.

```
Algorithm: CPEA (Classification procedure evaluation algorithm)
Input:   C (a set of characteristic descriptions)
Output: v (a value)
if any characteristic description in C covers an instance in NEG then set v to 0
else set v to the number of instances in POS covered by characteristic descriptions in C.
add $1 - \frac{1}{|C|}$ to v
```

As it does not relate to the difference between the two algorithms, the final step of ROA and TOA algorithms, refinement by general and domain specific rules, has been omitted.

The ROA and TOA algorithms were evaluated by application to a subset of the data and then evaluating the performance of the resulting classification procedure against the remnant of the data. In each test, instances were individually randomly selected for the training set with a 75% chance of inclusion. As both algorithms are sensitive to the order of instances in the training set, the data was randomly shuffled before each use. The algorithms were applied 100 times for each even setting of *width* from 0 to 20.

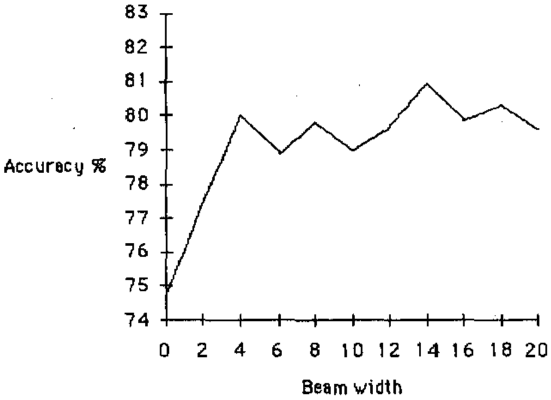
To generate a complete classification procedure, the respective algorithms were applied twice to each training set, once with *POS* set to contain those instances diagnosed with Immunoglobulin A Nephropathy disease and once with *POS* set to contain all other instances. The characteristic descriptions formed were expressed as production rules and sorted in order of the number of positive cases that they covered from highest to lowest. During evaluation, the first rule to apply to a case was selected.

The language used to express characteristic descriptions supported conjunctions of conditions each relating to a single attribute. For nominal attributes a condition tests whether the value for an attribute falls within a specified set of values. A missing value is treated as a distinct value. For ordinal attributes a condition either tests whether the value of an attribute falls within a specified range or requires that it is missing.

The Appendix shows the output of a run of the TOA algorithm with *width* set to 20.

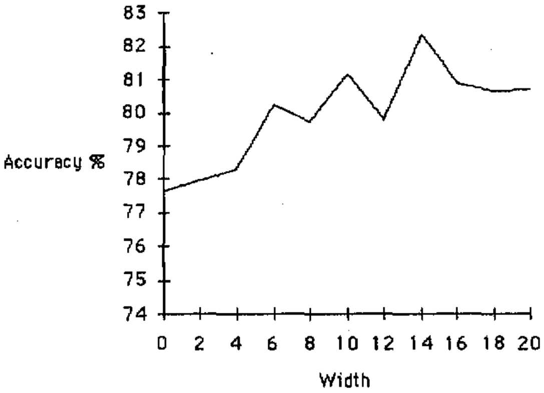
Figures 1 and 2 chart the accuracies of the two algorithms at successive settings of *width*. As can be seen, as the setting of width increases, both algorithms show an initial increase in accuracy which appears to rapidly level out. The average accuracy over all runs of ROA was 79.1% and of TOA was 79.9%.

Although these results suggest a modest advantage for TOA over ROA, this fails to take account of the fact that identical settings of *width* are not commensurate across the two algorithms. Rather than attempting to compare performance on the basis of performance for given setting of width, a more relevant measure is accuracy against computation time. Figure 3 compares the performance of the two algorithms on this basis. This comparison shows TOA consistently providing higher accuracy than ROA for equivalent expenditures of computation time.



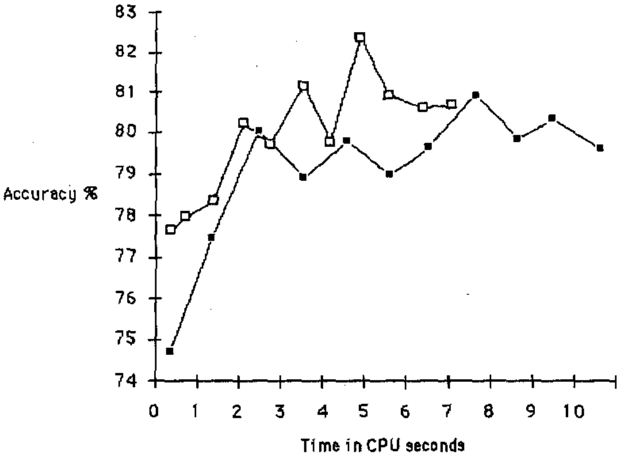
Each point represents the average accuracy over 100 runs at the specified setting of *width.*

**Figure 1: Rule optimisation accuracy**



Each point represents the average accuracy over 100 runs at the specified setting of *width*

**Figure 2: Theory optimisation accuracy**



Each point represents the average accuracy and average computation time in CPU seconds over 100 runs at a single setting of *width* [■ = ROA, □ = TOA]

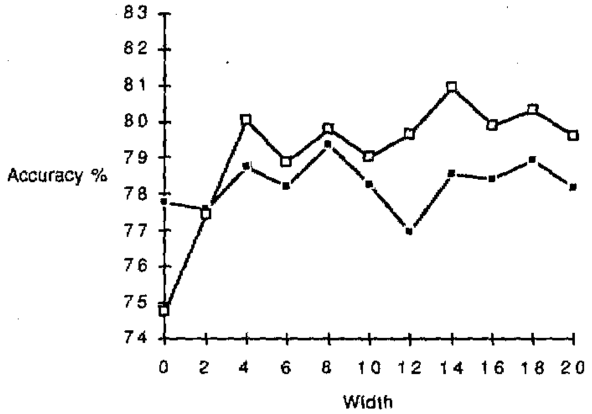
**Figure 3: Comparative performance of ROA and TOA plotting accuracy against computation time.**

One measure of the power of an induction algorithm is to compare its performance against chance performance (the accuracy obtained by making random classifications.) With two classes, random performance would he expected to provide 50% accuracy. Both algorithms perform considerably better then this.

Another interesting measure of the power of an induction algorithm is to compare its performance against the *Default Classification Strategy* (always classifying an instance as belonging to the most common class.) This extremely simple strategy gives an indication of the difficulty of obtaining high accuracies for a domain. For each of the experimental runs of the system, the accuracy obtained by the Default Classification Strategy was measured. Figure 4 compares the performance of the Default Classification Strategy and ROA and Figure 5 compares the performance of the Default Classification Strategy and TOA. These results show that both algorithms perform consistently better than the Default Classification Strategy with settings of *width* above two.

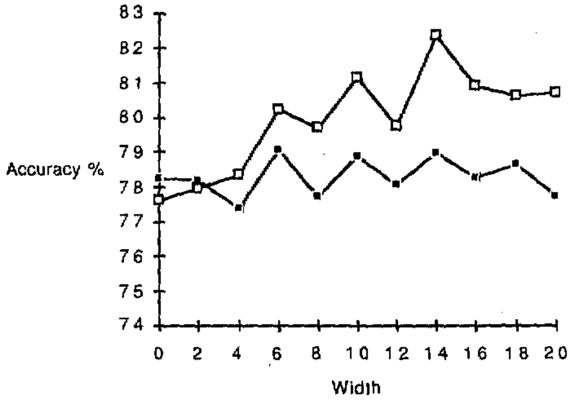
The accuracy of the Default Classification Strategy equals the proportion of the data withheld for evaluation which belongs to the most common class. The variations observable in Figures 4 and 5 in the performance of the strategy reflect random variations in the proportion of the most common class in the withheld data. Correspondences between variations in the proportion of the most common class in the withheld data and variations in performance of ROA and TOA that are apparent in Figures 4 and S suggest that this is one of the factors affecting performance of both ROA and TOA. This accounts to some extent for the ragged nature of the performance curves for the two algorithms observable in Figures 1 to 5.

A final comparison to draw between the two strategies is the number of characteristic descriptions that are produced. The evaluation criteria have been tailored to produce as few characteristic descriptions as possible on the assumption that this is a good measure of the simplicity of a classification procedure and that the simpler classification procedure is more likely to be accurate. Figure 6 plots the number of characteristic descriptions produced against computation time. As can be seen, TOA is considerably more successful than ROA in producing smaller sets of characteristic descriptions.



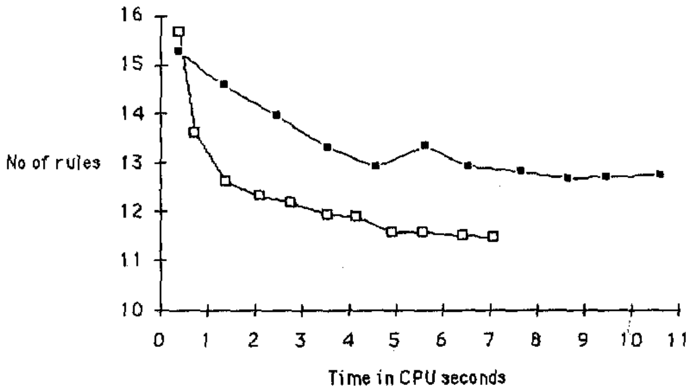
Each point represents the average accuracy over 100 runs at a single setting of *width*
[■ = Default, □ = ROA]

**Figure 4: Comparative performance of ROA and the Default Classification Strategy**



Each point represents the average accuracy over 100 runs at a single setting of *width*
[■ = Default, □ = TOA]

**Figure 5: Comparative performance of TOA and the Default Classification Strategy**



Each point plots the average number of characteristic descriptions produced against the average computation time in CPU seconds over 100 runs at a single setting of *width*
[■ = ROA, □ = TOA]

**Figure 6: Comparative performance of ROA and TOA plotting number of characteristic descriptions produced against computation time**

It is worthy of note that the large reduction in the number of characteristic descriptions produced by TOA has not produced such a dramatic improvement in performance. The average accuracy of all runs of TOA was only 0.7% greater than the average accuracy of all runs of ROA. By contrast, the average number of rules produced by all runs of TOA was 8.0% smaller than the average number of rules produced by all runs of ROA. Further, comparing runs that produced comparable numbers of rules (such as TOA with width two and ROA with width four), ROA produces higher accuracy. This suggests that although TOA is considerably better at developing sets of characteristic descriptions that maximize the evaluation criteria, those criteria do not provide an ideal measure of the value of a set of classification procedure. TOA can be expected to provide even greater improvements in accuracy over ROA when the evaluation criteria employed more adequately measure the relative worth of different classification procedures.

## 3 Conclusion

Previous approaches to the generation of classification procedures based on characteristic descriptions have utilized search heuristics that attempt to maximize the value of the entire classification procedure by finding a succession of characteristic descriptions each of which individually is of maximal value. This rule optimisation is achieved by generating multiple candidate rules and selecting the rule with the highest value.

This strategy has been contrasted with a strategy that generates and selects between multiple candidate classification procedures. The theory optimisation strategy has the advantage that the final object of interest, the classification procedure, is the direct object of evaluation. However, the possibility remains that the single candidate classification procedure generation algorithm employed will not produce sufficient high value classification procedures for the strategy to be successful.

This study has demonstrated that this possibility is not realized, at least for the domain of the diagnosis of Immunoglobulin A Nephropathy disease. For an extremely small set of data with high levels of missing values theory optimisation has out performed rule optimisation.

### Acknowledgements

I would like to thank John Agar and Geelong Hospital for providing the data used in this study.

## Bibliography

[1] Winston, P. H. **Learning structural descriptions from examples**. In P. H. Winston (Ed.), *The Psychology of Computer Vision*, McGraw Hill, New York, 1975, pp. 157-209.

[2] Quin J. R. **Induction of decision trees**. *Machine Learning*, 1986 **1**:81-106

[3] Quinlan, J. R. **Simplifying decision trees**. *IJMMS*, **27**:221-234(1987),

[4] Webb, G. **Techniques for efficient empirical induction**. *Proceedings of the Second Australian Joint Conference on Artificial intelligence*, Adelaide, 1988, pp. 218-232.

[5] Mitchell, T. M. **Version Spaces: A Candidate Elimination Approach to Concept Learning**, *PhD thesis*, Stanford University, 1979.

[6] Michalski, R. S. **Synthesis of optimal and quasi-optimal variable-valued logic formulas**. *Proceedings of the 1975 International Symposium on Multiple Valued Logic*, Bloomington, Ind., 1975, pp 76-87.

[7] Michalski, R. S. **A theory and methodology of inductive learning**. R. S. Michalski, ,J. G. Carbonell, & P. M. Mitchell (Eds.), *Machine Learning: An Artificial Intelligence Approach*, Springer-Verlag, Berlin, 1984, Pp. 83-129.

[8] Larson, J. B. **Inductive Inference in the Variable Valued Predicate Logic System VL21: Methodology and Computer Implementation**. *PhD thesis*, Computer Science Department, University of Illinois, Urbana, Ill, 1977.

[9] Plotkin, G. D. **A note on inductive generalization**. In B Meltzer & D. Michie (Eds) *Machine intelligence 5*, Edinburgh University Press, Edinburgh, 1970, pp. 153-163.

## Appendix

The following is output generated by the TOA algorithm with *width* set to 20. Each characteristic description is expressed as a production rule. The number of positive cases covered by each rule is listed in brackets at the end of the rule.

```
IF          2<=AGE<=77
            100<=SYST_BP<=230
            60<=DIAST_BP<=120
            DYSPNOEA is T
            800<=U.RBC's<=500000
            0.l0<=UTP<=23.13
            0.l0<=CRCL<=10.50
            30.00<=S.CR<=680.00
            3.0<=S.UREA<=30.6
            0.20<=C3<=1 96
            0.03<=C4<=l.21
            5.4<=HB<=16.6
            3.8<=WCC<=31.8
            I00<=PL'TS<=648
            5<=ESR<=142
            ANF is unknown or NEG or P0S
            ANTI-GBM is unknown or NEG or P0S
THEN DIAGNOSIS = other [45]
```

```
IF          15<=AGE<=72
            105<=SYST_BP<=210
            60<=DIST_BP<=115
            H'PTYSIS is F
            BOWEL_SX is F
            0<=U.RBC's<=33000
            0.43<=UTP<=16.25
            0.l8<=CRCL<=3.23
            70.00<=S.CR<=561.00
            3.2<= S.UREA<=43.8
            42<=TOT.PROT<=89
            28<=S.ALB<=52
            7.4<=HB<=20.4
            4.3<=WCC<=11.0
            140<=PL'TS<=280
            ANF is unknown or NEG or P0S
            ANTI-DNA is unknown
            ANTI-GBM Is unknown or NEG
THEN DIAGN0SIS = other [25]
```

```
IF          2<=AGE<=68
            90<=SYST_BP<=160
            50<=DIAST_BP<=100
            DYSPNOEA is F
            H'PTYSIS is F
            0<=U.RBC's<=500000
            0.20<=UTP<=11.00
            40.00<=S.CR<=112.00
            IGG is unknown
            IGM is unknown
            ANF is unknown or NEG or POS
            ANTI-GBM is unknown
THEN DIAGNOSIS = other [28]
```

```
IF          15<=AGE<=66
            110<=SYST_BP<=230
            65<=DIAST_BP<=160
            H'PTYSIS is F
            RASH is F
            WT_L0SS Is F
            1000<=U.RBC's<=640000
            0.10<=UTP<=6.79
            0.50<=CRCL<=2.76
            60.00<=S.CR<=1188.00
            3.5<= S.UREA<=18.9
            2.6<=IGA<=5.40
            11.7<=HB<=17.3
            3.7<=WCC<=9.8
            I66<=PL'TS<=299
            ANF is unknown or NEG or P0S
            ANTI-DNA is unknown
            ANTI-GBM is unknown
THEN DIAGN0SIS = IGA_NX[24]
```

```
IF              2<=AGE<=79
                100=SYST_BP<=260
                60<=DIAST_BP<=150
                ANOREXIA is T
                8000<=U.RBC’s<=500000
                0.62<=UTP<=22.69
                65.00<=S.CR<=1000.00
                3.1<= S.UREA<=78.6
                0<=TOT.PROT<=74
                0<=S.ALB<=51
                8.0<=HB<=15.1
                ANF is unknown or NEG or POS
                ANTI-GBM Is unknown or NEG or POS
THEN DIAGNOSIS = other [24]
```

```
IF              8<=AGE<=58
                120<=SYST_BP<=180
                65<=DIAST_BP<=100
                OEDEMA is F
                DYSPNOEA is F
                H’PTYSIS is F
                22000<=U.RBC’s<=500000
                0.15<=UTP<=1.47
                0.60<=CRCL<=3.58
                47.00<=S.CR<=138.00
                2.5<= S.UREA<=7.9
                7.90<=IGG<=14.4
                0.90<=IGA<=3.20
                0.34<=IGM<=2.60
                S.CHOL is unknown
                S.TRIGLY is unknown
                11.2<=HB<=16.3
                4.4<=WCC<=16.0
                174<=PL’TS<=550
                2<=ESR<=41
                ANF is unknown or NEG
                ANTI-GBM Is unknown or NEG
THEN DIAGN0SIS = other[13]
```

```
IF              4<=AGE<=78
                80=SYST_BP<=210
                50<=DIAST_BP<=120
                LOIN PAIN is F
                ANOREXIA is F
                WT_L0SS Is F
                0<=U.RBC’s<=500000
                2.44<=UTP<=28.97
                30.00<=S.CR<=938.00
                3.7<= S.UREA<=32.3
                9.5<=HB<=16.6
                4.6<=WCC<=13.6
                187<=PL’TS<=590
                ANF is unknown or NEG or POS
                ANTI-GBM Is unknown or NEG
THEN DIAGNOSIS = other [21]
```

```
IF              13<=AGE<=55
                110=SYST_BP<=180
                70<=DIAST_BP<=100
                HEADACHE is F
                RASH is F
                ANOREXIA is F
                BOWEL_SX is F
                112000<=U.RBC’s<=500000
                0.13<=UTP<=2.00
                0.85<=CRCL<=3.32
                52.00<=S.CR<=270.00
                3.9<= S.UREA<=12.0
                S.CHOL is unknown
                S.TRIGLY is unknown
                S.GLUCOSE is unknown
                11.0<=HB<=15.7
                5.6<=WCC<=8.8
                244<=PL’TS<=467
                3<=ESR<=58
                ANF is unknown or NEG or POS
                ANTI-GBM Is unknown
THEN DIAGN0SIS = IGA_NX[8]
```

```
IF              11<=AGE<=76
                105=SYST_BP<=200
                70<=DIAST_BP<=115
                H'PTYSIS is F
                BOWEL_SX is F
                5000<=U.RBC's<=500000
                0.31<=UTP<=2.56
                0.52<=CRCL<=2.12
                50.00<=S.CR<=200.00
                63<=TOT.PROT<=76
                29<=S.ALB<=47
                S.GLUCOSE is unknown
                11.3<=HB<=17.5
                7.0<=WCC<=12.5
                164<=PL'TS<=436
                1<=ESR<=118
                ANF is unknown or NEG
                ANTI-GBM Is unknown or NEG
THEN DIAGN0SIS = IGA_NX[17]
```

```
                6<=AGE<=72
                SEX is M
                90=SYST_BP<=200
                55<=DIAST_BP<=120
                RASH is F
                BOWEL_SX is F
                5000<=U.RBC's<=114000
                0.22<=UTP<=1.60
                50.00<=S.CR<=1083.00
                3.4<= S.UREA<=52.1
                ASOT is unknown
                S.CHOL is unknown
                S.TRIGLY is unknown
                S.GLUCOSE is unknown
                8.1<=HB<=13.5
                3.2<=WCC<=24.6
                100<=PL'TS<=433
                ANF is unknown or NEG
                ANTI-DNA is unknown
                ANTI-GBM Is unknown or NEG
THEN DIAGNOSIS = other [7]
```

```
IF              AGE=11
                SEX is M
                110=SYST_BP<=240
                55<=DIAST_BP<=130
                HEADACHE is T
                DYSPNOEA is F
                H'PTYSIS is F
                U.R.T.I. is T
                ANOREXIA is T
                BOWEL_SX is T
                300000<=U.RBC's<=500000
                0.30<=UTP<=0.50
                0.27<=CRCL<=1.47
                100.00<=S.CR<=320.00
                17.8<=S.UREA<=24.0
                8.80<=IGG<=22.0
                1.70<=IGA<=2.85
                1.10<=IGM<=1.22
                0.20<=C3<=2.08
                0.18<=C4<=0.60
                S.CHOL is unknown
                S.TRIGLY is unknown
                S.GLUCOSE is unknown
                10.5<=HB<=11.0
                7.4<=WCC<=8.4
                460<=PL'TS<=608
                36<=ESR<=55
                ANF is NEG
                ANTI-DNA Is unknown
                ANTI-GBM Is unknown
THEN DIAGNOSIS = other [2]
```