

獲得知識を援用した MTS 法による肝疾患判定法

1 序論

本研究では、獲得知識を援用した MTS 法による肝疾患の自動判定を行った。知識の獲得にはラフ集合を用い、判定では、血液検査項目を用い MTS 法で肝疾患を病態(肝疾患なし・肝癌・急性肝炎・慢性肝炎・肝硬変)に 5 分類した。

肝疾患の病態での分類では、分類が多種多様にわたり、関連する血液検査項目が多く複雑で、さらに確実な診断を下すためには、画像・肝生検等の検査が必要となる。そこで、血液検査結果を多次元ベクトルと考へて距離尺度を導入する方法が行われてきた[1]。しかし、これらは純粹に統計学の立場に基づいたもので、そこで扱うデータも単に数値としての意味しか持ち合わせていない。

医学知見や医師の経験などを何らかの方法で診断システムに組み込むことができれば、診断精度は飛躍的に向上することが期待できる。

2 肝疾患について

肝疾患の兆候は、黄疸、肝汁うっ滞、腹水、肝機能障害、肝不全、肝性昏睡(脳症)、門脈圧亢進症、肝腎症候群がある。

肝疾患に関連する血液検査項目として、TbIL, Alp, γ -GTP, LDH, ChE, GPT, GOT, PLt, Alb, AFP の 10 項目を本研究では使用した。ただし、肝疾患の判定に有用な γ -グロブリン、ZTT(硫酸亜鉛混濁試験)、T-Chol(総コレステロール)等も用いたかったが、それらの実際の血液検査データを入手できなかったため、本研究では使用していない。また、GOT/GPT 比も診断には欠かせない指標となる。

3 ラフ集合

3.1 ラフ集合とは

ラフ集合とは、1982 年にパブラックによって提唱された対象の集合をうまく特定できる範囲で情報を荒く(ラフに)することで、対象の集合の程よい記述を求める手法である。対象を識別するのに必要最低限の属性の集合や対象が所属するクラスを識別する簡潔なルールなどを導き出す方法を与える。ラフ集

合の中では、決定ルールの考え方が最も多く使われており、そこで求められた複数の決定ルール条件部のどれがどれだけ結論に寄与しているかを示す指標 C.I.(Covering Index)が分析結果の考察に有効なものさしとなる。

3.2 ラフ集合による知識獲得

縮約と決定ルール条件部を出力では、市販のラフ集合解析プログラムを使用した。

縮約では今回利用したデータを判別するための必要最小限の項目は Alb 以外全ての検査項目であることがわかった。確かに Alb は本来、慢性肝炎や肝硬変等で低値を取るはずだが、今回利用したデータで Alb の値はほとんど低値を見られなかったため、この縮約は正しい。

さらに、決定ルール条件部で肝疾患の病態 5 つについて数学的に特徴を見出すことができた。ただし、属性値数の上限があるため、今回のような連続したデータ等は属性値数を新たに定める必要があり、その設定により結果が左右されてしまう。また、サンプル数の上限が 100 個であるため、今回のように曖昧なデータが多い場合、縮約でもこれ以上検査項目を絞ることができず、C.I.が低い結果となってしまう。

4 MTS 法

4.1 MTS 法とは

MTS 法(Mahalanobis Taguchi System)とは、データの偏りを考慮したマハラノビス距離を利用したタグチメソッドのことである。つまり、TMS 法では、ある単位空間に対して求めたマハラノビス距離をその単位空間に属するかどうか判定する。マハラノビス距離が小さいほど、単位空間に属すると考え、閾値を決めて判定する。

4.2 MTS 法による肝疾患判定法

ここで使用した血液検査結果は、TbIL, Alp, γ -GTP, LDH, ChE, GPT, GOT, PLt, Alb, AFP の 10 項目である。データは、単位空間の平均値 a_j と標準偏差 σ_j 、相関行列の逆行列 $\{b\}$ を用いて、式(1)で基準化し、式(2)でマハラノビス距離 D_i を求めた。

$$u_{ij} = \frac{x_{ij} - a_j}{\sigma_j} \quad (1)$$

$$D_i^2 = (u_{i1} \dots u_{ij}) \begin{pmatrix} r_{11} & \dots \\ & \dots \\ & \dots & r_{jj} \end{pmatrix} \begin{pmatrix} u_{i1} \\ \vdots \\ u_{ij} \end{pmatrix} \quad (2)$$

閾値は ROC 曲線を描いて決め、表 1 のような結果を得た。

表 1. 従来の MTS 法の結果

従来の MTS 法	TP(%)	FP(%)
肝疾患なし	100	1
肝臓	60	25
急性肝炎	65	54
慢性肝炎	86	48
肝硬変	84	22

この結果は、肝疾患なしについては特に良い結果が得られているが、他は FP 率が高くなってしまった。これは、血液検査結果が、データのばらつきだけでなく、検査項目間の比に注目する場合や、複数の空間に属したりどの空間にも属さなかったりという曖昧なデータを含んでいる可能性があるといった特徴を持っているからである。

5 提案手法

MTS 法による肝疾患判定を行う際に、ラフ集合で得られた知識を援用し判定する。

5.1 MTS 法からラフ集合へ移行するルール

MTS法での判定で差が微妙だと思われるデータについてラフ集合を適用することとした。単位空間が基準となることを前提に、閾値の上下(単位空間と同じ幅)を差が微妙なデータとしてラフ集合へ移行して再判定を行い、それより下は単位空間に属する、それより上は単位空間に属しないとした。ラフ集合へ移行するルールは式(3)である。 A_{max} , A_{min} は単位空間の最大値と最小値、 a_n は病態 n におけるMTS法での閾値である。

$$a_n - \frac{|A_{max} - A_{min}|}{2} < \log D_i < a_n + \frac{|A_{max} - A_{min}|}{2} \quad (3)$$

5.2 MTS 法とラフ集合の併合

MTS 法とラフ集合の併合で得られた結果は表 2 であった。

提案手法により TP 率は上がったとは言えないが、肝臓や急性肝炎では、TP 率がそのままに FP 率を下げる事ができた。

表 2. 提案手法の結果

獲得知識を援用した MTS 法	TP(%)	FP(%)
肝疾患なし	84	0
肝臓	53	8
急性肝炎	65	19
慢性肝炎	73	40
肝硬変	58	8

ROC 曲線で見ても、特に急性肝炎で精度が上がったことがわかる(図 1)。

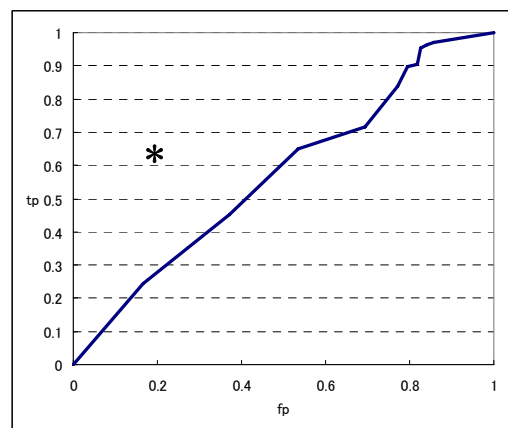


図 1. ROC 曲線による比較(急性肝炎)

6 結論

本研究では、獲得知識を援用した MTS 法による肝疾患判定法について研究を行ったが、獲得知識の援用ではラフ集合を用いて、MTS 法だけでは判定不可能だった曖昧なデータを判別することができた。

課題は、さらに診断精度を上げるため、MTS 法での閾値、MTS 法からラフ集合へ移行するルールなどについて検討し、ラフ集合解析プログラムの制約にとらわれないことで、これらの手法を拡張し、予後予測や知識抽出にも応用させたい。

参考文献

[1] 神崎訓枝, 金川明弘: 血液検査項目による Mahalanobis 距離肝疾患判定法, 第 27 回医療情報学連合大会論文集, pp.515-516, 2007.