

平成 15 年度弁理士試験論文式筆記試験問題

[情報工学]

1 . データマイニングに関する事柄につき、以下の問いに答えよ。

【 26 点】

(1) 表 1 は、あるスーパーマーケットにおける顧客 10 人の商品購入リストを示している。この表から、「相関ルール A B」の確信度 $\text{conf}(A \rightarrow B)$ およびサポート $\text{support}(A \rightarrow B)$ 、ならびに「相関ルール C E」の確信度 $\text{conf}(C \rightarrow E)$ およびサポート $\text{support}(C \rightarrow E)$ を、それぞれ計算せよ。ただし、「相関ルール X Y」の確信度は X を含むトランザクションのうち Y も含むものの割合で表され、サポートは全トランザクションのうち X と Y とを両方含むものの割合で表される。ここでトランザクションとは、表 1 商品購入リストにおいて個々の顧客に対応する 1 行分のデータをいう。すなわち、表 1 のリストは 10 個のトランザクションから成る。

表 1

顧客	購入商品
1	A, B, D, E
2	B
3	A, D
4	C, D, E
5	A, B
6	B, D
7	A, D, E, F
8	B, E
9	A, B, F
10	A, B

(2) 発見された相関ルールの価値を測る上で、「確信度」および「サポート」が指標としてどのような意義を有するか、考察せよ。

(3) 多次元データのクラスタリングを行う代表的な技法を一つ挙げ、そのアルゴリズムを説明せよ。

(4) データマイニングにおける「例外データ(outlier)」の意義について、
() 例外データの影響を排除することが重要である場合と () 例外データそのものを検出することが重要である場合とに分けて、具体例を挙げて論ぜよ。

2 . 以下の語句の内から 4 つを選んで説明せよ。

【 2 4 点】

(1) コンピュータ・ビジュアライゼーション

(2) GPS(Global Positioning System)

(3) マルチモーダル・インタフェース

(4) 遺伝的アルゴリズム

(5) SQL(Structured Query Language)

(6) レイトレーシング法 (コンピュータ・グラフィックス技法の一つ)

論点 [情報工学]

- 1 . データマイニングに関する以下の諸点についての理解を問う。
 - (1) 与えられた定義どおりに信頼度およびサポートを計算できるか
 - (2) 信頼度およびサポートの指標としての役割
 - (3) 基本的なクラスタリング技法のアルゴリズム
 - (4) データマイニングにおける例外データの意義

- 2 . 最新のトピックスや基本的な学術用語についての一般的な知識と理解を問う。