

研究テーマ：データマイニングの調査

067番 余川 洋史

1、研究課題

人工頭脳のひとつであるデータマイニングについて調査し、その概要を理解する。また現在データマイニングを用いた実用例や研究も調査し、さらにどのように発展させていけるかを自分なりに考察する。

2、調査結果

データマイニングとは、データウェアハウスなどに蓄積された、一見因果関係もないように見える膨大なデータから、共通したパターンやデータ間の因果関係をもとに、未知の知見や規則性を発見するプロセスのことである。

データマイニングを行う流れは、大まかに「データの収集」、「データの選択」、「データの前処理」、「データの変換」、「学習・解析」、「予測・評価」の6つのステップに別けることが出来る。また場合によっては、この「学習・解析」のステップだけをデータマイニングと呼ぶこともある。

この「学習・解析」のときに用いる手法には決定木、回帰木、相関分析、回帰分析、クラスター分析、ニューラルネットワークなどといったものがあり、行う解析に適した解析手法を選択しなければならない。

データマイニングの応用例としては、医療や健康に関わるものを調査した。

調査したもののひとつは「小児脳腫瘍とデータマイニング」についてである。これは現在医師の主観に頼っている脳腫瘍の分類を、データマイニングによって正確に分類し、脳腫瘍に対する最適な治療法の発見を目指すものである。

二つ目に調査したのは「臨床研究とデータマイニング」についてである。これは、多くの患者から得られたデータをデータマイニングで分析することで、医療現場の医師たちが、より効果の見込める治療方針を選択できるようにするものである。

最後に調査したのは「インターネットを利用した健康データマイニング」についてである。これは、日常生活習慣データと健康データをインターネット経由でデータを蓄積し、グラフ表示や、データマイニングによって発見した生活習慣データと健康データの間規則性を、ユーザにそれらを通知するというものである。

3、考察

最後の調査した「インターネットを利用した健康データマイニング」を、複数の人に対して行い、その結果をさらにデータマイニングを行うことで、今後起こりうる健康状態も予想し、より優れた健康管理ができるのではないだろうか。