

ニューラルネットによる防災マッチングのための Web データ分類

山口大学大学院 学生会員 ○小俣尚泰
 (株) 栗本鐵工所 関根聡一
 山口大学工学部 フェロー会員 宮本文穂
 山口大学工学部 高城和志

1. はじめに

近年、駿河トラフ、南海トラフを震源とする東海・東南海・南海地域での地震の発生による大規模かつ広範囲の災害の発生が懸念されており、過去の経験を基にして各機関において対策が急ぎ進められている¹⁾。そこで、本研究では災害の現場で発生するニーズに適した人材・物資・機材・情報が迅速に手配できるようなマッチングシステムの構築を目指している。本システムでは、シーズデータベースから最適なシーズを探索し、手配する。マッチングシステムにおいて重要な部分となるシーズデータベースは、構築・保守において、多大な労力を要するため、それらのコストを削減することが望まれる。本研究では、この問題に対し World Wide Web (以下 Web) に注目し、Web からの情報を利用してシーズデータベースを構築することを試みた。

2. マッチングシステムと Web データ分類

マッチングシステムとは一定の規則に従って効率的にかつ迅速に需要 (needs) と供給 (seeds) を一致 (matching) させる仕組みの総称である。本研究では、このマッチングシステムの仕組みを利用した災害復興活動支援を行うシステムの構築を目指している。マッチングシステムを導入すると図 1 のように、ニーズとシーズを把握し、種々の情報をデータ

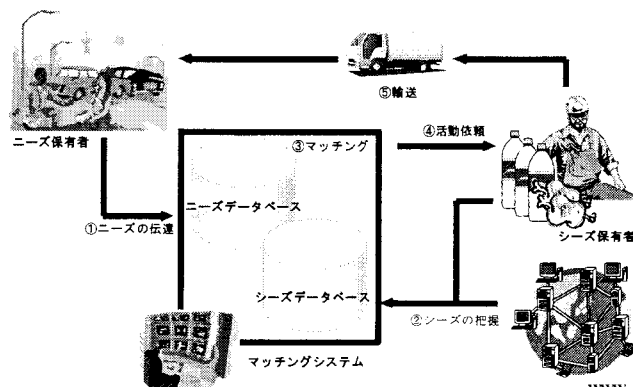


図1 システム導入後のワークフロー

シーズ名	シーズ保有者名	保有数	調達可能数	提供可能数	貸与可能数
シーズ1	山口大学	1	10	1	1
デスクトップパソコン	山口大学	12	12	12	12
デスクトップパソコン	山口大学	30	12	12	12
デスクトップパソコン	山口大学	30	30	12	12

検索結果一覧

http://www.123.com/content/default.asp?cid=123&pid=456
 http://www.456.com/pc/index.htm
 http://www.789.com/pc/index.htm
 http://www.101.com/Computers/Internet/Hardware/Personal Computers/IBM/Computers/FMV/
 http://www.123.com/
 http://www.456.com/pc/index.htm
 http://www.789.com/pc/index.htm
 http://www.101.com/

図2 マッチング結果例

ベース化する。そして、システム内部で災害の状況などの各種情報と照らし合わせながら対応させ適材適所にシーズを送り込む。図 2 にマッチングシステムの画面例を示す。後述する Web データ分類を用いることにより、Web 上のデータを結果提示の補助として使うことによって、情報が見つからないという状況をなくすることが可能となる。

ここで、マッチングを行うための情報が見つからないという状況を極力排除するために、シーズデータベースは多くの情報を用意していなければならない。そのため、データベース構築時に情報の自動獲得を行い、適切に分類され蓄積されれば、この問題は解決できると考えた。

3. ニューラルネットワーク²⁾

ニューラルネットワークは生物の神経回路における神経細胞をモデル化することにより実現された学習機械である。その一手法である誤差逆伝播法はネットワークの出力と教師信号との差を小さくするように「重み」と「しきい値」を更新していく学習方法である。以下にそのアルゴリズムを示す。

① 初期の「重み」と「しきい値」を乱数により決める。

② 入力と期待出力の提示

入力 $x_0, x_1, \dots, x_{n-1}$ と教師信号 $t_0, t_1, \dots, t_{m-1}$ をネットワークに与える。ここで n および m はそれぞれ入力

および出力のユニット数である。 ω_0 をしきい値、 χ_0 を常に 1 とする。

③ ネットワークの出力

入力層から出力層に向けて各ユニットの出力を、式(1)より計算する

$$o_{pj} = f\left[\sum_{i=0}^{n-1} \omega_{ij} \chi_i\right] \quad (1)$$

④ 重みの更新

式(2)を用いて出力層から入力層に向けて重みの更新を層ごとに行う。

$$\omega_{ij}(t+1) = \omega_{ij}(t) + \eta \delta_{pj} o_{pi} \quad (2)$$

ここで、 $\omega_{ij}(t)$ は t 回目の学習におけるユニット i から j への重み、 η は重みの変化量を調節する学習率、 δ_{pj} はパターン p に対するユニット j の誤差項を表す。

δ_{pj} は出力ユニットに対しては式(3)より、

$$\delta_{pj} = k \cdot o_{pj}(1 - o_{pj})(t_{pj} - o_{pj}) \quad (3)$$

隠れ層に対しては式(4)より、それぞれ計算される。

$$\delta_{pj} = k \cdot o_{pj}(1 - o_{pj}) \sum_k \delta_{pk} \omega_{jk} \quad (4)$$

式(3),(4)の和はユニット j が属する層の上の層に属するユニット k すべてにわたってとるものとする。

⑤ 繰り返し

ネットワークの出力が教師信号に近づくまで繰り返す。

4. Web データ分類実験

ニューラルネットワークを用いた Web データの分類の評価実験を行った。識別器に入力するベクトルの基底数を 100 個、1,000 個、10,000 個の 3 パターンと定め、中間ユニット数を 1, 2, 4, 8 の 4 パターンとして実験を行った。検証方法には Leave-one-out 法を用いた。

実験の結果として、入力ベクトルの基底数とネットワークの誤識別率の関係を図 3 に示す。図 3 には、昨年度に行った SVM(Support Vector Machine)³⁾を用いた同様の実験⁴⁾の結果を含めている。入力ベクトルの基底数と適合率の関係を図 4 に示す。中間ユニット数 8 の場合に良い誤識別率が得られていることがわかる。中間ユニット数が少ない時に誤識別率が高くなっているのは学習に失敗していることによる。また、基底数は誤識別率に大きく作用することではなく、計算時間に大きく作用している。結果

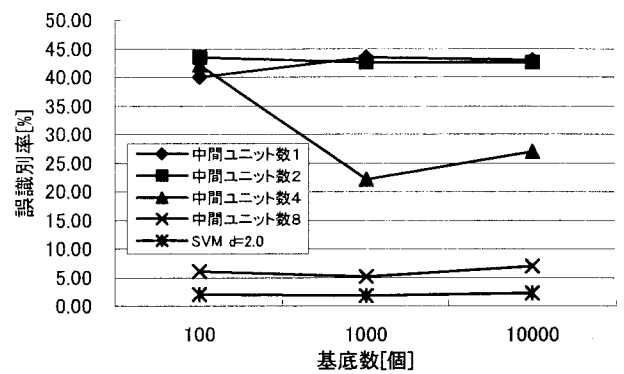


図 3 基底数と誤識別率の関係

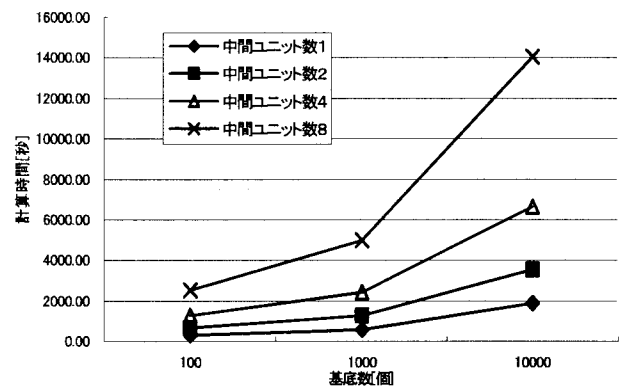


図 4 基底数と計算時間の関係

より、中間ユニット数を増やすと計算時間にとって有利ではなくなるため、ユニット数は 8 が適当であることがわかる。

5. まとめ

本研究によって得られた結果を以下に示す。

- ① 防災マッチングシステムのための Web データ分類システムを構築した。
- ② 誤識別率については適切な中間ユニット数を指定することにより、SVM とほぼ同等の分類性能を持つことが確認できた。

参考文献

- 1) 内閣府防災部門：わが国の災害対策，内閣府政策統括官（防災担当），2002.3.
- 2) 熊沢逸夫：学習とニューラルネットワーク，森北出版，1998.7.
- 3) 前田英作：痛快！サポートベクトルマシン，情報処理学会誌，42，pp.676-683，2001.7.
- 4) 小俣尚泰ら：救援活動支援および防災対策のためのマッチングシステムの提案，第 56 回平成 16 年度土木学会中国支部研究発表会発表概要集，p.p.565-566，2004.5.