

ニューラルネットワークを利用したデータマイニングの土木工学への適用

武蔵工業大学 学生会員 佐藤 大介
 武蔵工業大学 正会員 星谷 勝
 (株)地崎工業 正会員 須藤 敦史

1. はじめに^{1),2)}

本研究は、ニューラルネットワークを用いて、東京湾で観測された赤潮発生に関する水質データ、宍道湖・中海(鳥取県)で観測された水質データ、北海道のトンネルにおける覆工コンクリートのひび割れ観測データのそれぞれに対してデータマイニングを行うものである。図-1 に示すように、3 事例においてデータの特徴を考慮した解析を行い、観測項目間の関連性を明らかにしている。

2. ニューラルネットワークによる解析

解析事例(1)は、観測データ間の関連性の発見・把握を目標とし、階層型ニューラルネットワークを採用している。ネットワークの重みの大きさから観測データ間の関連性の強さを算出し、東京湾における赤潮発生のメカニズムを明らかにしている。解析事例(2)は、宍道湖・中海の水質データに対して、階層型ネットワークを発展させた擬似的な相互結合型ネットワークを形成し、観測項目間の因果関係を明らかにしている。解析事例(3)は、寒冷地におけるトンネル覆工コンクリートのひび割れ発生要因の分析を行う予定である。

なお、観測値の単位やディメンジョンに統一性がないことから、観測されたデータをそのままの形式で扱うと、簡潔なネットワークが構築できないため、対象とするすべてのデータを平均値以上・未満で二分し、ブール属性値(0-1 データ)に変換して扱っている。その際、ニューラルネットワークの学習法は、成長側抑制学習を採用しているが、それは、成長側抑制学習を用いるとネットワークの重みの成長に抑制をかけて、重要な結合だけが生き残るように調節するため、ニューロン間の結合が明確になるという理由からである。

3. 解析事例(1)

この解析では、出力項目が唯一であり、ニューラルネットワークを利用する際に最も適したデータ形態といえる。「水温」、「透明度」、「NH₄-N」は平均値未満が、「COD」、「T-P」、「T-N」、「sal.」、「chl-a」は平均値以上が赤潮発生と関連しているというルールが得られた(表-3)。さらに、結合の強さ(太さ)から相関の度合いが評価でき、解析によって得られた赤潮発生に関するルールを関連性の大きい順に以下に示す。

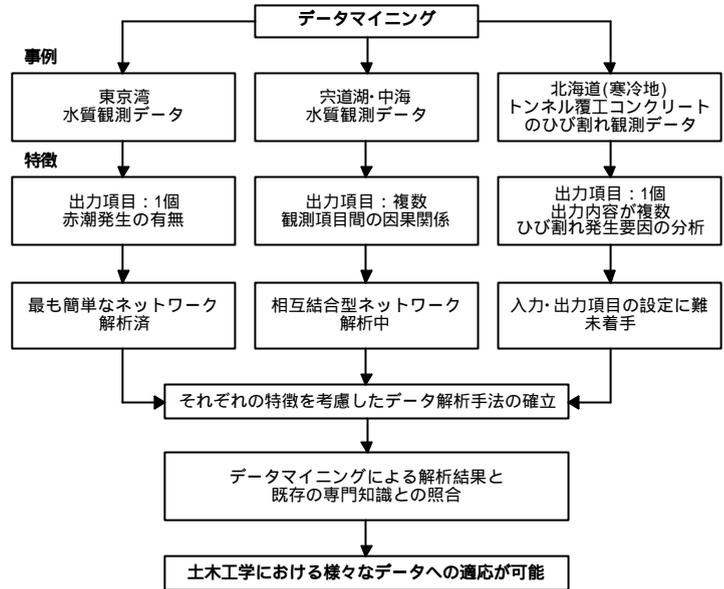


図-1 本研究のフロー

表-1 解析事例(1)観測項目

東京湾の水質データ		説明
A	気温	()
B	水温	()
C	透明度	(m)
D	pH	-
E	COD	化学的酸素要求量(mg/l)
F	DO	溶存酸素量(mg/l)
G	T-P	全リン(mg/l)
H	PO ₄ -P	リン酸態リン(mg/l)
I	T-N	全窒素(mg/l)
J	NH ₄ -N	アンモニア態窒素(mg/l)
K	NO ₂ -N	亜硝酸態窒素(mg/l)
L	NO ₃ -N	硝酸態窒素(mg/l)
M	sal.	塩分(mg/l)
N	chl-a	クロロフィル a(mg/l)
O	赤潮	赤潮発生の有無

表-2 解析事例(2)観測項目

宍道湖・中海の水質データ		説明
A	trans.	透明度(m)
B	w-temp.	水温()
C	pH	-
D	sal.	塩分濃度(%)
E	DO	溶存酸素量(ppm)
F	turb.	濁度
G	SS	懸濁物質(mg/l)
H	chl-a	クロロフィル a(ppb)
I	COD	化学的酸素要求量(ppb)
J	T-P	全リン(ppb)
K	PO ₄ -P	リン酸態リン(ppb)
L	T-N	全窒素(ppb)
M	NO ₂ -N	亜硝酸態窒素(ppb)
N	NO ₃ -N	硝酸態窒素(ppb)
O	NH ₄ -N	アンモニア態窒素(ppb)

Key word : データマイニング, ニューラルネットワーク, 相関関係解析

連絡先 : 〒158-8557 東京都世田谷区玉堤 1-28-1 03-3703-3111

第1ルール：水温：小，COD：大，chl-a：大

第2ルール：透明度：小，NH₄-N：小

第3ルール：T-P：大，T-N：大，sal.：大

また「pH」に関しては平均値以上・未満の両方の関連性が抽出されているが、これは他の項目との相互関係が影響していると考えられるため、今回の分析ではルールから除外している。

4. 解析事例(2)

2つの湖で観測されたデータの中から、すべての観測項目を出力に1回ずつ設定し、残りの14項目を入力に設定し、解析を行った結果を表-4に示す。水深別・位置別解析では、湖全体の解析で得られた関連性とほぼ同様の結果が得られているの

が分かる。また、深さ別では水面付近でCOD - T-P、湖底付近でpH - sal.に対して相関を有している結果が抽出された。場所別では宍道湖でw-temp. - NO₃-N、sal. - NO₃-N、T-N - NO₃-N、中海でsal. - T-Nに対して新しい相関が抽出された。湖全体の解析で強い相関があった8項目を用いて図-2のような結合ネットワークを作成した。次に、8項目の以外の項目について、再度、解析結果の見直しを行った(例えば、trans. - SS)と

ころ、強い相関があった8項目と比較すると弱いが、図-2で矢印が結ばれていない4箇所も相関を有しており、観測項目間で相関の正・負の関係が一致していることから、図-3のような擬似的な相互結合型ネットワークを構築することができた。

5. 結論と今後の展望

土木工学における、ニューラルネットワークを用いたデータマイニングの適用について検討を行った。各観測項目を平均値以上・未満のブール属性値に変換し、成長抑制学習を用いて重みの小さくなった結合を削除することにより、ネットワークの簡素化が実現し、ルールの抽出が容易になった。また、今後の課題として、解析事例(3)におけるデータ解析手法の確立が挙げられる。ひび割れ発生要因の分析を目標としているが、ひび割れに関するデータは、発生の有無、発生場所、形状など多くの内容を含むことから、出力項目の設定が困難であり、工夫が必要であると考えられる。また、現段階までは、データマイニングの基本概念に基づき、専門知識は考慮せずに観測データの関連性を分析することに主眼を置いてきたが、今後は、既存の専門知識とデータマイニングにより得られた知識を比較検討する予定である。

<参考文献>

- 1) 須藤敦史・佐藤大介・星谷勝「ニューラルネットワークを用いたデータマイニングの環境問題への適応について」第6回システム最適化に関するシンポジウム講演論文集 pp.23-28,1999.12.
- 2) NEUROSIM™/L light,富士通,1996.

表-3 解析事例(1)の結果

属性	結合の強さ	結合状態	ルール
B: 水温	強	(+) (-)	「小」 「赤潮あり」 「大」 「赤潮なし」
C: 透明度	中		
J: NH ₄ -N	中		
E: COD	強	(-) (-)	「大」 「赤潮あり」 「小」 「赤潮なし」
G: T-P	弱		
I: T-N	弱		
M: sal.	弱		
N: chl-a	強		
D: pH	弱	(+) (-)	「小」 「赤潮あり」 「大」 「赤潮なし」
	弱	(-) (-)	「大」 「赤潮あり」 「小」 「赤潮なし」

表-4 解析事例(2)の結果

湖全体	深さ別		場所別	
	水面付近	湖底付近	宍道湖(1,2,3)	中海(3,8,11,24)
trans. - turb. w-temp. - DO DO - PO ₄ -P Turb. - SS chl-a - COD T-P - PO ₄ -P T-N - NH ₄ -N NO ₃ -N - NO ₂ -N	trans. - turb. w-temp. - DO turb. - SS COD - T-P T-P - PO ₄ -P NO ₃ -N - NO ₂ -N	w-temp. - DO ph - sal DO - PO ₄ -P Turb. - SS chl-a - COD T-P - PO ₄ -P T-N - NH ₄ -N NO ₃ -N - NO ₂ -N	trans. - turb. w-temp. - DO w-temp. - NO ₃ -N sal. - NO ₃ -N turb. - SS T-P - PO ₄ -P T-N - NO ₃ -N T-N - NH ₄ -N NO ₃ -N - NO ₂ -N	w-temp. - DO sal. - T-N DO - PO ₄ -P chl-a - COD T-P - PO ₄ -P NO ₃ -N - NO ₂ -N

は湖全体と比較して新たに現れた相関

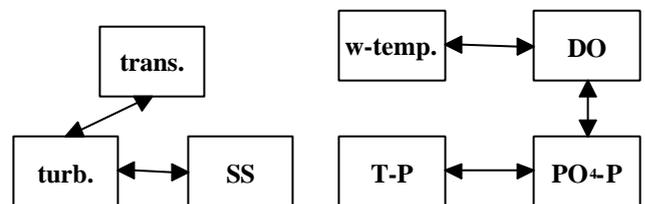


図-2 結合ネットワーク

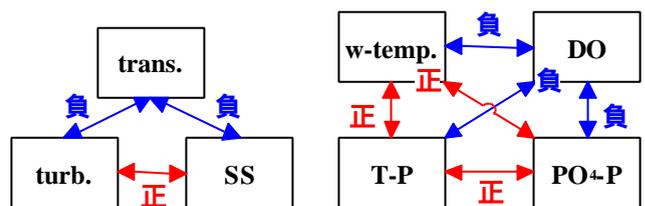


図-3 擬似相互結合型ネットワーク