

AI学習による解析事例



中日本建設コンサルタント(株) 水環境技術本部 技師長 中根 進

1 はじめに

既報¹⁾では、降雨時に流下する下水温とAI学習(ニューラルネットワーク法。以下、nnetと略記する)により降雨時における晴天水温を推定し、降雨日の水温と降雨量の推移を図示した。

AI学習(nnet)を利用した解析は、水温のみならず、流量、管内水位にも適用でき、報告²⁾した。

本報告では、雨天時浸入水量の推定を具体例にAI学習(nnet)を使った手法を説明する。また、AI学習(nnet)で脱水ケーキの含水率を回帰した例を示し、AI学習(nnet、ランダムフォレスト、ベクターサポートマシン)を使って管路の劣化程度と管路台帳項目を分析した例を示す。

nnetによりコンクリート管の常時浸入水要因の分析³⁾やAI学習の一つであるOne class SVM(ワンクラス サポートベクターマシン)を使って電動機の動作音の異常検知に利用する試み⁴⁾も行ったが、これらは参考文献を参照されたい。

2 AI学習

AIは、「認識・判断」「認識・理解」「予測、推

論」「計画・最適化」などを実行する人間の知能を人工的に模倣するための総合的な概念技術である。この総合技術は単一の要素技術ではなく、多くの要素技術を含んでおり、これらの技術を適切に選択し、目的に対して適用することにより、その能力を発揮させるものである⁵⁾。

人工的に知能を模倣するためには、法則やルール、事象を学習する必要がある、その学習技術を機械学習といい、多くのアルゴリズム(学習方法)がある。アルゴリズムには、2変量またはそれ以上の変量の関係から解析する「回帰」、ランダムフォレストなどの「木」や「ニューラルネットワーク」という手法など多くの技術がある。代表的な機械学習アルゴリズムを図-1に示す。

3 ニューラルネットワーク法の概要

3.1 ニューラルネットワークのアルゴリズム

AIの学習技術の一つであるnnetは、人間の脳のしくみ(ニューロン間のあらゆる相互接続)から考えられたもので、脳機能の特性のいくつかをコンピュータ上で表現するために作られた数学モデルと言われている。

このアルゴリズムを模式的に図-2に示す。

入力層の各ユニット（ニューロン）にデータ x_i を入力し、入力値に重み v を掛けた代数和 $u_j = \sum_i (v_{ij} \cdot x_i)$ を求め、中間層に入力し、中間層で出力層への出力値を計算する。中間層の各ユニットの z_j に重み ω を掛けた代数和 $a_k = \sum_j (\omega_{jk} \cdot z_j)$ を求め、出力層で $y_k = f(a_k)$ を計算し、実測あるいは観測流量である教師データと比較する。

中間層や出力層での出力には、(1)式などが使わ

れる。筆者はシグモイド関数を使った。

$$z = f(u) = \begin{cases} 1 & u \geq 0 \text{ の時} \\ 0 & u < 0 \text{ の時} \end{cases} \dots\dots(1)a$$

$$z = f(u) = u \dots\dots\dots(1)b$$

$$z = f(u) = \frac{1}{1 + e^{-u}} \dots\dots\dots(1)c$$

：シグモイド関数
(ロジスティック関数)

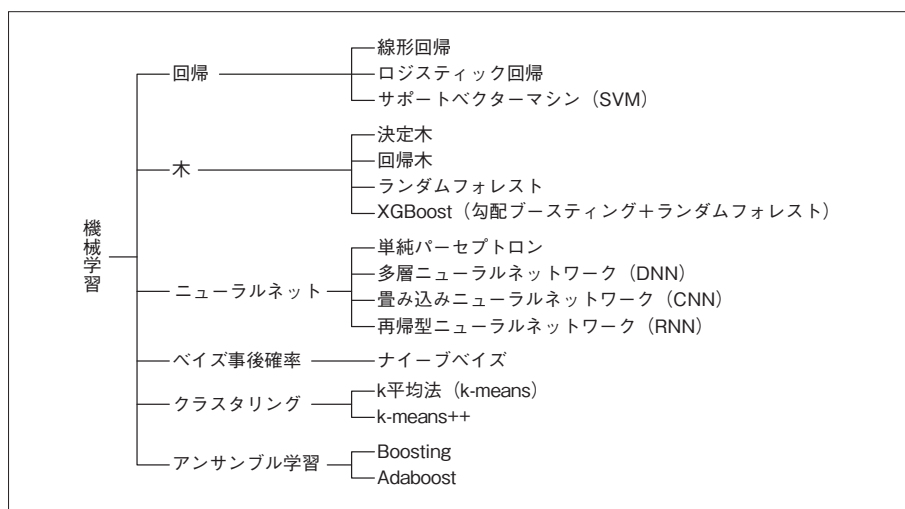


図-1 代表的な機械学習の学習技術

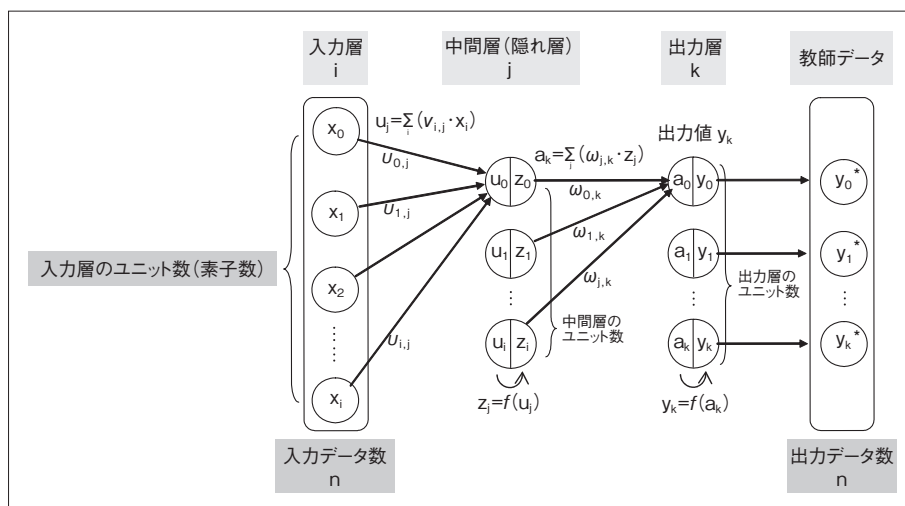


図-2 ニューラルネットワークのアルゴリズム

3.2 ニューラルネット法による 雨天時浸入水量解析モデル

管きょ流量調査や処理場流入流量の時系列データを解析する場合のモデルを説明する。

筆者らの一人が文献³⁾で利用したnnetのアルゴリズムを図-3に示す。このnnetは入力層-中間層-出力層の3層と中間層を複数持つ深層からなるフィードフォワード型ニューラルネットワークで、学習方法には誤差逆伝搬法（バックプロパゲーション法）を用いている。図-4は中間層を2層設定した図となっている。それぞれユニット（ニューロン）は複数設定できる。今までに解析した結果では、中間層が1層でも複数層でも解析結果が変わるものではない。

入力層には、カレンダー情報などの説明変数を割り当てるため、雨天時浸入水量の推定では表-1に示す91個のユニット（ニューロン）を設けた。出力層は、ある時刻の計測データのみに対応するため、1個のユニット（ニューロン）を設ける。また、中間層にはn個のユニット（ニューロン）を設ける。ユニット数により出力値と教師データ間の相関が変わるため、相関が高くなるユニット数を設定する。今までの解析事例からは中間層のユニットは3～5個で十分であった。ここで使う量的データは、多変量解析で実施するように解析前にすべて正規化（規準化）する。質的データはダミー変数として扱い0と1で表わす。

表-1の量的データ48個は、降雨後の降雨影響時間であり、下水処理場では48時間程度あると言われている。小面積の流入区域を持つ流量計測点では、少ない時間となる。

筆者は、FIRフィルターを使い降雨影響時間を設定している。一例を図-5に示す。

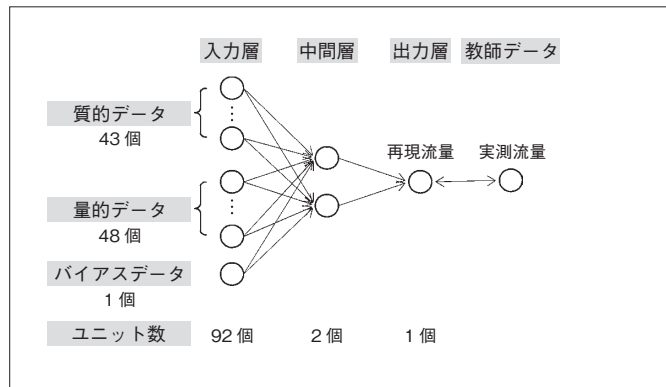


図-3 中間層1層のニューラルネットワーク

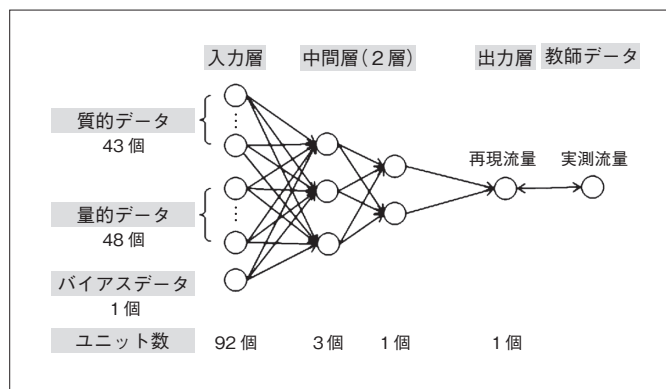
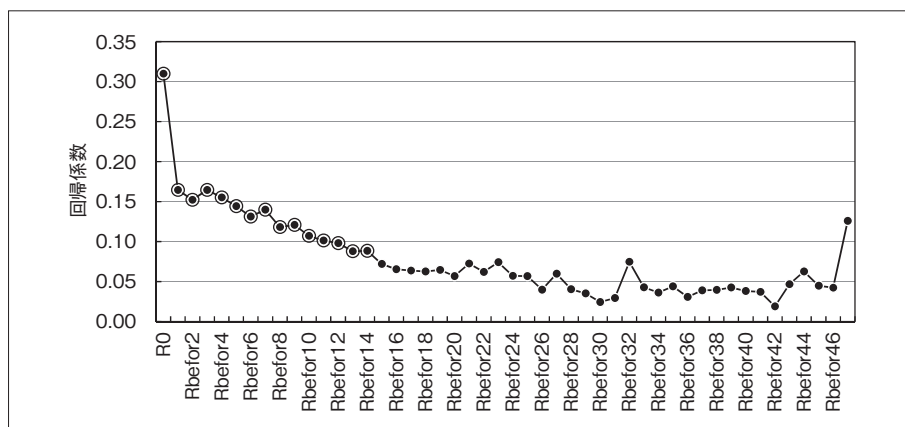


図-4 中間層2層のニューラルネットワーク

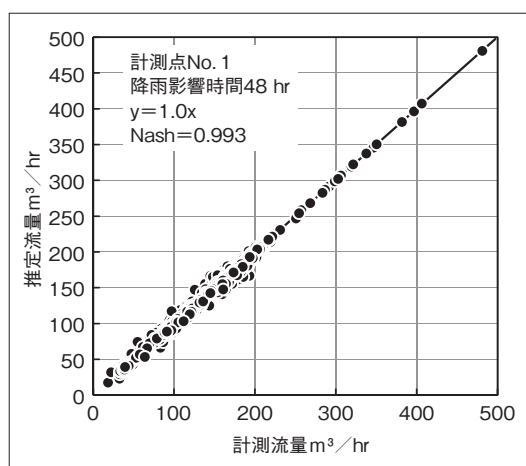
表-1 流量解析のデータ(説明変数)構成

データ項目	データ構成
教師データ	・ 計測流量
目的変数	・ 推定流量
説明変数	
入力層データ：時刻カテゴリ（1時間間隔）	
質的データ	・ 時刻カテゴリ：24個※1
	・ 曜日カテゴリ：7個※2
	・ 月別カテゴリ：12個※3
	小計 43個
量的データ	・ 1時間降雨量データ：48個※4
	小計 48個
合計 91個	

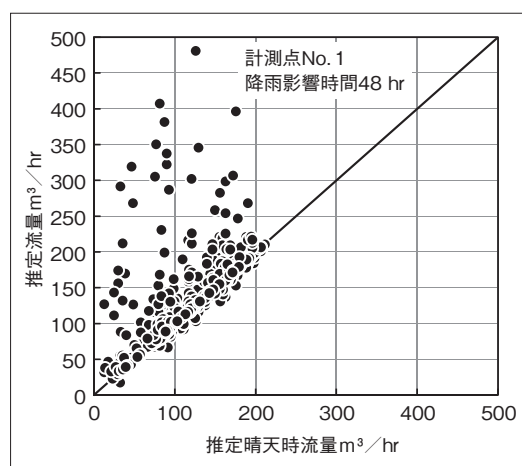
- ※1 時刻カテゴリ：1日（24時間）の流量の周期性を表す
 ※2 曜日カテゴリ：1週間の生活活動の変動を表す
 ※3 月別カテゴリ：調査期間3ヵ月の季節変動を表す
 ※4 対象とする計測流量の降雨の影響がある時刻（2日分）まで遡る降雨量の1時間間隔データ



図－5 降雨影響時間の回帰係数



図－6 計測流量と推定流量の関係 (No. 1)



図－7 推定晴天時流量と推定流量の関係 (No. 1)

雨天時浸入水量に及ぼす降雨影響時間はその回帰係数の適合具合を示す p 値によって判断する。 p 値 <0.025 を有意として図中に○印を付す。降雨影響時間は14時間となり、降雨終了後14時間まで降雨の影響が続くと推定できる。

4 ニューラルネット法による雨天時浸入水量解析事例

nnetにより上記表－1のデータ構成に示す説明変数を使って目的変数とする計測流量を推定(以下、推定流量という)する。

流量の時系列データを設定した説明変数を使って回帰し、計測流量を再現する。得られた説

明変数に対する重みを用いて降雨をなし(0)として再計算することにより雨天時の晴天時水量を推定する。計測流量と推定晴天時水量の差が雨天時浸入水量となる。

一例として流量調査箇所No.1(管径600mm、流量計[超音波式水中流速計+水位計])の計測流量とニューラルネット法で解析した推定流量の当てはまり具合を図－6に示す。計測流量と推定流量は45度線上に乗っている。また、推定晴天時流量と推定流量の関係は図－7となり、45度線より上が雨天時浸入水量となる。

図－6にあるNash(Nash-Sutcliffe)係数は、河川の洪水流量解析に際し、洪水流量と推定洪水

流量の回帰具合を表すために多く使われる。なお、Nash係数は、1.0を最大とし1.0に近いほどモデルの精度は良いとされ、0.7以上でモデルの再現性が高いとされている。筆者の感覚では0.9以上でないと再現性が高くないように思われる。

計測点No. 1のある降雨日の時間変動を図-8に示す。図中に回帰した流入水量を(—)で示し、流入水量の計測値を打点(●印)で示す。打点した計測流量と回帰した推定線はよく一致している。降雨時における降雨ない水量(推定晴天時流量)を(—)で示す。

5 ニューラルネットワークによる 脱水機操作因子と脱水ケーキ含水率

脱水ケーキ含水率を目的変数、脱水機の操作因子を説明変数としてnnetで解析した事例を示す。

脱水ケーキ含水率(目的変数)に及ぼすと思わ

れる、以下に示す脱水機投入汚泥の性状や操作因子(ろ過速度、薬注率)などの管理情報(説明変数)を正規化してnnetで解析する。管理情報からは①気温、雨量、②PAC注入量、③濃縮槽pH、④生污泥固形物割合、消化污泥固形物割合、余剰污泥固形物割合、⑤高分子凝集剤13種類、⑥高分子凝集剤注入率、⑦脱水機ろ過速度、⑧カレンダー情報(月、曜日)を抽出し説明変数とする。下水処理場において約4年間、脱水機稼働日1日1回定時に測定した脱水ケーキ含水率の推移を図-9(—)に示す。

解析した結果は図-9に併せて示す。推定した脱水ケーキ含水率(—)は、ばらつきのある計測した脱水ケーキ含水率(—)をうまく推定していないが、平均的な値を示した。上記の説明変数以外の因子を探して加える必要があると考えられる。

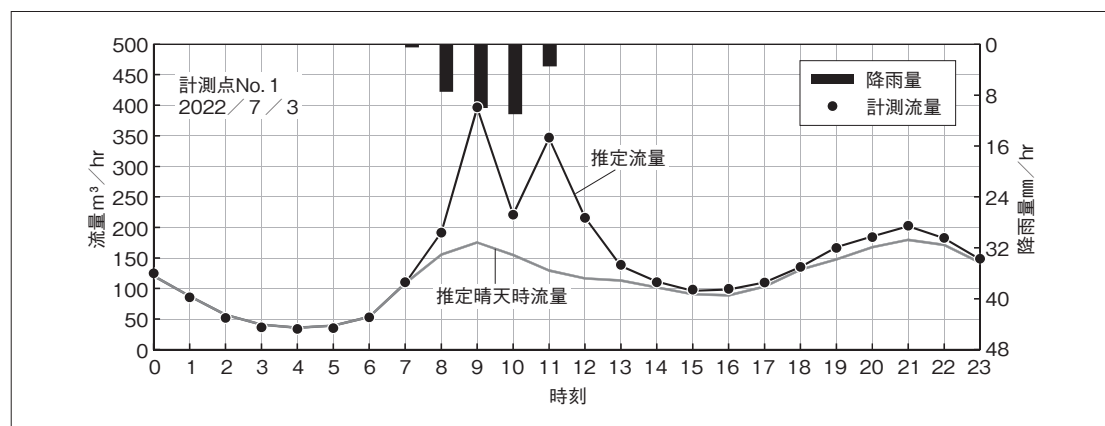


図-8 降雨日の計測流量、推定流量と推定晴天時流量の推移

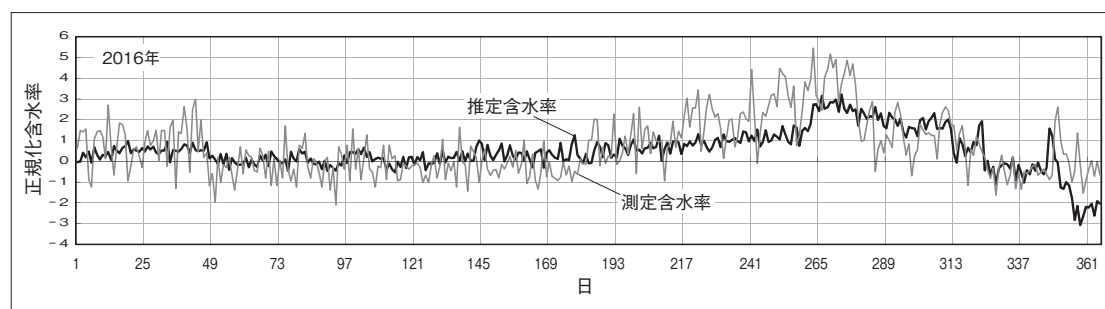


図-9 脱水ケーキ含水率と推定脱水ケーキ含水率の推移

6 管路の劣化程度と 管路台帳項目の分析

(1) ニューラルネットワークによる管路台帳項目 と緊急度で表わした劣化程度

コンクリート管のTVカメラ調査診断結果である緊急度（Ⅰ、Ⅱ、Ⅲ、Ⅳ）に影響している管路台帳項目（説明変数）をnnetで推定する。

推定した管路台帳項目（説明変数）を使い、自治体の持つ管きょの緊急度を推定し、緊急度の高い管路台帳項目を示す管きょから点検・調査を実施するなどの机上のスクリーニングとして利用する。

国土技術政策総合研究所「下水道管きょ劣化データベース Ver.2」にある情報を使い、コンクリート管の緊急度（目的変数）と説明変数にする下記管路台帳項目の関係を解析する。

管渠劣化データベースにある管路台帳項目（説明変数）は、次の管種から土被りまでの8項目とする。

- 管種 : ヒューム管
- 管径 : 250mm～1,000mm
- 路線延長 : 5.36m～64.03m
- 管本数 : 3本～30本
- 取付け管本数 : 1本～8本
- 経過年数 : 7年～35年
- 歩車道区分 : 歩道、車道
- 土被り : 0.8m～63.3m

データベース中に説明変数と緊急度の記載があるスパン数を415スパン抽出し、このデータを解析する。隠れ層（中間層）のニューロン数を4とし、出力層に緊急度Ⅱ、Ⅲ、Ⅳ（健全）を配置して解析した。

解析した結果を表－2に示す。各スパンの緊急度に対して、出力は、各緊急度の確率値として示される。スパン番号1の出力結果は、緊急度Ⅱ＝0.012、緊急度Ⅲ＝0.820、緊急度Ⅳ（健全）＝0.168であり、確率の最も高い緊急度Ⅲを推定緊急度とする。表－2をとりまとめて表－3に示す。

表－2 調査診断緊急度と推定緊急度

推定判定：0 不一致／1 一致

スパン 番号	調査診断 緊急度	推定緊急度の確率			推定 緊急度	推定 判定
		Ⅱ	Ⅲ	Ⅳ		
1	Ⅲ	0.01157	0.82014	1.68E-01	Ⅲ	1
2	Ⅳ	0.04116	0.39599	5.63E-01	Ⅳ	1
3	Ⅳ	0.025315	0.333857	6.41E-01	Ⅳ	1
4	Ⅳ	0.015027	0.345599	6.39E-01	Ⅳ	1
5	Ⅳ	0.000876	0.446202	5.53E-01	Ⅳ	1
6	Ⅳ	5.35E-09	0.340125	6.60E-01	Ⅳ	1
7	Ⅳ	0.014906	0.44914	5.36E-01	Ⅳ	1
8	Ⅲ	0.022364	0.310713	6.67E-01	Ⅳ	0
9	Ⅳ	1.57E-07	0.79117	2.09E-01	Ⅲ	0
409	Ⅲ	0.109479	0.47431	4.16E-01	Ⅲ	1
410	Ⅲ	0.149534	0.68735	1.63E-01	Ⅲ	1
411	Ⅲ	5.85E-08	0.498566	5.01E-01	Ⅳ	0
412	Ⅲ	3.68E-06	0.714269	2.86E-01	Ⅲ	1
413	Ⅲ	2.7E-07	0.720947	2.79E-01	Ⅲ	1
414	Ⅲ	3.14E-08	0.606716	3.93E-01	Ⅲ	1
415	Ⅳ	8.79E-06	0.641503	3.58E-01	Ⅲ	0

表-3 ニューラルネットワークによる緊急度推定結果

総合的中率 0.600

	推定結果		
	II	III	IV
II	0	0	0
III	14	167	54
IV	14	84	82
診断結果	28	251	136
的中率	0.000	0.665	0.603

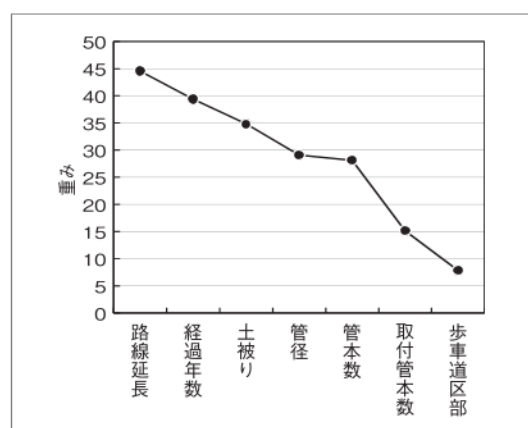


図-10 説明変数の重み

調査診断の緊急度と推定した緊急度は、全415スパンに対して60.0%しか一致していない。調査時における緊急度Ⅱのスパン数は28あるが、Ⅲ、Ⅳと推定してしまったスパン数はそれぞれ14あり、まったく推定できていない。

(2) ランダムフォレストによる管路台帳項目と緊急度の推定

学習用のデータをランダムにサンプリングして多数の決定木を作成し、作成した決定木をもとに多数決で結果を決める方法である。精度、汎用性が高く扱いやすい分析手法と言われている。

「6(1)」と同じデータを用い、ヒューマンのTVカメラ調査診断結果である緊急度（Ⅰ、Ⅱ、Ⅲ、Ⅳ）に及ぼす影響のある要因（説明変数）をランダムフォレストで推定する。

ランダムフォレストによる解析結果は表-4と

表-4 ランダムフォレストによる緊急度推定結果

総合的中率 0.986

	推定結果		
	II	III	IV
II	26	1	0
III	2	250	3
IV	0	0	133
診断結果	28	251	136
的中率	0.929	0.996	0.978

表-5 SVMによる緊急度推定結果

総合的中率 0.781

	推定結果		
	II	III	IV
II	3	0	0
III	21	230	45
IV	4	21	91
診断結果	28	251	136
的中率	0.107	0.916	0.669

なる。調査時における緊急度Ⅱのスパン数は28であり、ランダムフォレストによる推定では、緊急度Ⅱと推定したスパン数が26で、誤って緊急度Ⅲとしたのが2スパンあり、緊急度Ⅱの的中率は92.9%であった。同様に緊急度Ⅲは、ランダムフォレストで251スパン中250スパンを推定しており、的中率は99.6%になった。ランダムフォレストでは緊急度Ⅱ、Ⅲ、Ⅳのすべての的中率は98.6%となり、nnetの60.0%より高い的中率となった。ランダムフォレストによる緊急度に及ぼす説明変数の重みは図-10となり、スパン長の重みが大となり、続いて経過年数、土被りの重みが大となった。

(3) SVMによる管路台帳項目と緊急度の推定

SVM (Support Vector Machine) は「教師あり学習」に分類される。線形、非線形の識別関数があり現在知られている多くの学習モデルのなかでは、最も優れた識別能力があるとされている。SVMを使って「6(1)」と同じデータを用い、解析する。解析結果は表-5となる。緊急度Ⅱ、Ⅲ、

IVを含めた全体の的中率は78.1%であり、緊急度Ⅱの的中率は小さく、緊急度Ⅲの的中率は91.6%であった。緊急度Ⅲ、Ⅳの的中率が高くても学習結果の利用方法はあまりない。

7 まとめ

管路の台帳項目と緊急度の分析にはnnet法やSVM法より、ランダムフォレストのほうが優位であった。

〈参 考 文 献〉

- 1) 中根進：成分分解ならびに AI 学習による下水水温解析手法、月刊下水道、2025 年 1 月号、Vol.48、No. 1
- 2) 中根進ら：流量・水温法データの AI 機械学習による雨天時浸入水量解析の研究、下水道協会誌、2021 年 10 月号、VOL.58、No.708
- 3) 中根進：AI による污水管きよの常時浸入水に及ぼす影響要因の分析、環境浄化技術、2021 年 7・8 月号、Vol.20、No. 4
- 4) 中根進ら：動作音の計測・解析による異常検知の基礎的研究、N-11-24、第 60 回下水道研究発表会
- 5) NEC が考える「AI」とは：<https://jpn.nec.com/ai/relate/ai.html>、2025.01.08 確認