

令和 5 年 6 月 13 日現在

機関番号：37111

研究種目：基盤研究(C)（一般）

研究期間：2020～2022

課題番号：20K04691

研究課題名（和文）情報バイアスを有する地盤情報群を用いる層区分AI推定と精度評価

研究課題名（英文）A Stratigraphic Classification Estimation Method by Neural Networks for Geotechnical Information with bias and its Precision Evaluation

研究代表者

村上 哲（MURAKAMI, Satoshi）

福岡大学・工学部・教授

研究者番号：10261744

交付決定額（研究期間全体）：（直接経費） 1,900,000円

研究成果の概要（和文）：地盤情報DBに適用する人工知能を用いた地層区分推定方法を確立することを目的し、地盤情報は、調査地点、深度、そこで現れる土質など偏り（情報バイアス）を持つため、BD群の特徴を踏まえた上で、データを平均化して偏りを小さくし、推定位置に近いBDの情報ほど誤差が小さくなることを期待した地層区分推定のためのニューラルネットワーク（NN）の新しい全体誤差関数の考え方とその機械学習の方法を提案した。提案手法を福岡平野の地盤に適用し、その有効性を示した。また、機械学習に用いる地盤情報の統計値に基づく地層区分推定結果の精度評価手法を提案しその有効性を示した。

研究成果の学術的意義や社会的意義

地盤工学分野における人工知能技術の利用は、小田らによる圧密定数や強度定数の空間地盤物性値をニューラル・ネットワークにより推定するものなどがある。その他、地すべりなどの空間情報やグラフィカルな情報から画像解析、パターン認識より事象を抽出する問題への適用が多いものの、本研究で扱う層区分推定にAIを用いる研究は内外ともに例が無く、また、AIの推定値を求めるための重み係数を決定する評価関数について、地盤情報が情報バイアスある情報として捉えて再定義した研究事例は存在しない。地盤工学分野におけるAIの新しい利用法、適用範囲を広げることが期待できる学術的・社会的に貢献できる成果であると考えられる。

研究成果の概要（英文）：The purpose of this study is to establish a strata classification estimation method using artificial intelligence that can be applied to geotechnical information DB. We proposed a new concept of the overall error function of a neural network (NN) for strata classification estimation and a machine learning method for the function. The proposed method was applied to the ground in the Fukuoka Plain, and its effectiveness was demonstrated. We also proposed a method for evaluating the accuracy of the results of stratigraphic classification estimation based on the statistical values of the geotechnical information used in the machine learning, and demonstrated its effectiveness.

研究分野：地盤工学

キーワード：人工知能 機械学習 地盤情報DB 精度評価 地層区分 全体誤差関数 情報バイアス

様式 C-19、F-19-1、Z-19（共通）

1. 研究開始当初の背景

地盤の層区分は、建設計画・設計・施工、災害の復旧・対策立案において、工法の決定や災害のメカニズム及び対策の提案などの根拠の1つとして利用されている。高度な知識を有する技術者が実施・監修したものであれば、間違えないとしても良いかもしれないが、そうでない場合もありうるであろう。また、層区分された結果の確からしさがはっきりできない現状では、そこに潜在するリスクを事前に察知することが出来ないのではないかと考える。

この問いに答えるために、人工知能(AI)を用いた層区分推定的手法を確立する。特に、利用する地盤情報が深度、本数、出現する土質など偏りある情報（情報バイアス）である点を考慮した評価関数を提案し、その有効性・妥当性を実データによる数値実験により多面的に明らかにする。さらに、AI推定の特徴を活かし精度評価指標の提案を行う。

2. 研究の目的

地盤情報 DB に適用する人工知能を用いた地層区分推定方法を確立することを目的し、地盤情報は、調査地点、深度、そこで現れる土質など偏り（情報バイアス）を持つため、BD群の特徴を踏まえた上で、データを平均化して偏りを小さくし、推定位置に近いBDの情報ほど誤差が小さくなることを期待した地層区分推定のためのニューラルネットワーク(NN)の新しい全体誤差関数の考え方とその機械学習の方法を提案する。提案手法を福岡平野の地盤に適用し、その有効性を示す。さらに、機械学習に用いる地盤情報の統計値に基づく地層区分推定結果の精度評価手法を提案しその有効性を示す。以上の結果を踏まえて、情報バイアスを有する地盤情報群を用いる層区分AI推定と精度評価の提案を行う。

3. 研究の方法

人工知能(AI)のコア部には汎用的な階層型ニューラル・ネットワーク(NN)構造を利用し、偏りのある地盤情報群を用いる層区分推定を前提としたNNの評価関数を新たに提案し、本手法の有効性と妥当性を明らかにすることを目的とする。

具体的な適用性を示すために、福岡平野を対象とし研究を進める。福岡平野の第四紀系層序は上部より、埋土や盛土などの人工土層、完新世の中洲層、更新世の荒江層、博多粘土層、金武礫層に区分されている。この地域に存在する地盤情報約3万本を研究に利用する。地盤情報には層区分は未記録であるので、同図に示した福岡都市高速道路沿線の地盤情報に対し、各対象層の区分を行い、推定結果の正解値（学習データ）として用いることとする。

AI推定では、学習過程において、学習データと推定結果の残差の関数として定義される評価関数を最小化させることでモデルパラメータ（結合係数）を決定することが一般的であるが、本研究では、評価関数として次の3つのモデルを考える。

- ① 学習データと推定結果の残差二乗和を最小化させる（モデルⅠ）
- ② 層ごとに算出した残差二乗和の単純平均を最小化させる（モデルⅡ）
- ③ 地盤の空間変動から得られる距離相関を用い、層ごとに算出した残差二乗和の加重平均を最小化させる（モデルⅢ）

①は標準的なAIツールで利用される手法であり、取得したデータを分け隔てなく学習時の判断材料とする。しかし、地盤調査深度は一様ではなく、調査深度が増加するに従いデータ数は稀薄になる傾向となる。技術者は、このようなデータの偏りを認識した上で、各層ごとにその特徴を理解し層区分する。②、③はこのプロセスを評価関数に組み込んだモデルである。②と③の違いは、ある柱状図が与えられた場合、その周辺の地層の状況を見ながら順次決定していくか否かの違いであり、③はこの判断を評価関数に含んだモデルである。

本研究では、この①、②、③のモデルで層区分推定を行った場合の長所と短所を実地盤の推定結果により明らかにしていく。特に、①、②、③の順に、従来技術者が行ってきたプロセスを踏襲させることが層区分推定においてどの程度有効なのかを示す。加えて、評価関数の特性を生かし、地層区分AI推定の精度評価手法を提案し、その有効性を模擬地盤への適用により明らかにする。

4. 研究成果

(1) 情報バイアスを考慮した地層区分AI推定の定式化

今、地層区分された P 個のBD群があり、対象地盤において現れる層の総数が K である場合を考える。これらのBD内の地盤情報（深度、土質、 N 値など）を用いて、NNにより地層区分する場合を考える。

図-1に示すBD番号 p に着目する。層番号 k に該当するデータ個数を $N_{p,k}$ と表す。この地盤情報が入力値となり、入力値から中間層を経て出力値が算出されることとなる。入力値の個数を I とし、BD番号 p 、層番号 k 、 n 番目の学習データに対する i 番目の入力値を $z_{(p,k,n),i}$ と表す。この入力値を用いてNN解析を実施すると出力値が求められる。出力値は層区分数 K 個あり、 \bar{k} 番目の出力値を $x_{(p,k,n),\bar{k}}$ と表す。

BD 番号 p 、層番号 k 、 n 番目の学習データに対する誤差関数 $E_{(p,k,n)}$ について、 $x_{(p,k,n),\bar{k}}$ に対応する教師値 $o_{(p,k,n),\bar{k}}$ を用いて、次のように定義する。

$$E_{(p,k,n)} = \frac{1}{K} \sum_{k=1}^K (x_{(p,k,n),\bar{k}} - o_{(p,k,n),\bar{k}})^2 \quad (1)$$

一般に用いられている NN (本文では ONN と呼ぶ) では、全学習データに対する誤差関数の総和 (本研究では全体誤差関数 E と呼ぶ) を考え、これを最小化することで良い推定結果を得る NN モデルを決定する。この時、全体誤差関数は次式で定義される。

$$E = \frac{1}{N} \sum_{p=1}^P \sum_{k=1}^K \sum_{n=1}^{N_{p,k}} E_{(p,k,n)} \quad (2)$$

ONN 法は、データを分け隔てなく学習時の判断材料とする。しかし、BD の場合、調査深度は一樣ではなく、深度が増加するに従いデータ数は希薄になる傾向となる。技術者はこのようなデータの偏りを認識した上で、各層が持つ特徴を理解する。その上で、周辺の地層の状況を見ながら地層区分を決定する。そこで、ONN 法とは異なる全体誤差関数を改めて定義することにより、技術者の作法を取り入れることを考える。

今、推定したい BD と学習データの BD の距離に関する重み係数 $r_{p,k}$ を導入し、BD 内の地層ごとのデータの偏りを無くすように平均化して算出する次の全体誤差関数 E を定義する。

$$E = \frac{1}{K} \sum_{k=1}^K \sum_{p=1}^P \frac{r_{p,k}}{N_{p,k}} \sum_{n=1}^{N_{p,k}} E_{(p,k,n)} \quad (3)$$

ここで、距離に関する重み係数 $r_{p,k}$ の層ごとの総和は 1 となるように決定する。よりに応じた重み係数を用いる方法を D-Layer NN 法 (DLNN 法) と呼び、等しい重み係数とした場合を単純 Layer NN 法 (SLNN 法) と呼ぶことにする。SLNN 法は層ごとの個数に差がある教師データを平均化することで、その偏りを無くす工夫をしたものであり、DLNN 法はさらに推定地点に近い教師データとの差を小さくするものである。このように全体誤差関数を最小化させることにより、データを平均化して偏りを小さくし、推定位置に近い BD の情報ほど誤差が小さくなるような NN モデルとなることを期待する。

NN における機械学習では全体誤差関数 E を最小化させるように実行される。L2 正則化はこの関数 E に結合係数($w_i; i = 1 \sim n$)の 2 乗和に係数を乗じた次の関数値を最小化させることによつて過学習を回避させる方法である。

$$E + \frac{\lambda}{2} \sum_{i=0}^n w_i^2 \rightarrow \text{Min.} \quad (4)$$

ここで、 λ は正則化パラメータである。この方法を用いると、 λ が未知数となるため、機械学習時にこの値を決定する過程が含まれる。本研究では、まず、学習用データを教師データと検定データに分け、仮定した λ による教師データを用いた機械学習結果より得られる検定データの推定値と検定データの実測値の対比により λ を決定する。次に、決定した λ を用いて、DLNN 法で用いる IDW パラメータも、いくつか仮定した IDW パラメータによる教師データを用いた機械学習結果より得られる検定データの推定値と検定データの実測値の対比によりもっとも良いものを選定する。

以上のように、地盤情報の情報バイアスを考慮した地層区分 AI 推定の定式化と機械学習に用いる最小化手法を提案した。学習用データを用いた検討結果から、最も良いと思われる L2 正則化パラメータと IDW パラメータを用いて、検証用データの実測値と推定結果との対比により有効性について検討する。

(2) 情報バイアスを考慮した地層区分 AI 推定法の有効性の検討

① 対象地域の概要と学習用データ、検証用データの作成

対象地域は、福岡平野である。BD は九州地盤情報共有 DB(2005, 2012, 2018)を用い、図-2 に示した連続した地盤情報が存在する福岡都市高速千鳥橋 JCT(a)から月隈 JCT(b)の沿線の BD を用いた。DB から抽出した 125 本の BD について、産業技術総合研究所提供の「ボーリング柱状図解析システム ver. 2.1」

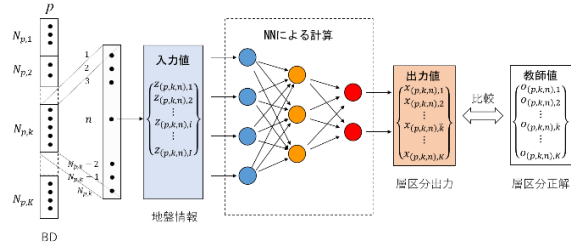


図-1 p 番目 BD の NN 計算と教師データの説明図

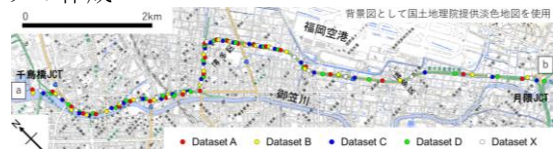


図-2 対象地域・断面線・BD の分布図

を用いて当該地域における地層区分を行った。地層区分を図-3に示す。125本のデータを5分割し、4つを学習用データ(Dataset A, B, C, D)とし、残り1つを検証用データ(Dataset X)とした。学習用と検証用のデータの傾向を図-4と5に示す。この図より、各層が占める割合に違いは見られないものの、中洲層、荒江層のデータ数が多く、その大半を占め、金武礫層は全体の1~1.5割程度と極めて少ない傾向にある。

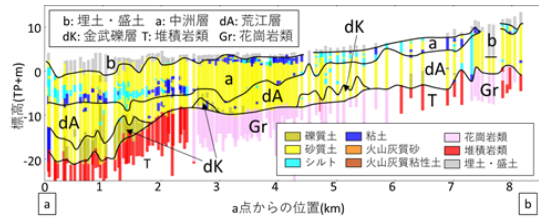


図-3 地層区分結果

② 情報バイアスを考慮した地層区分 AI 推定法の有効性の検討結果

Dataset A, B, C, Dの4組の学習用データを用いて、3組を教師データとし、残り1組を検定用データとして、ONN法とSLNN法により λ を決定する。 $\lambda = 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}$ を与えた結果、ONN法とSLNN法のそれぞれについて、教師データによる正解率と検定用データの正解率の差が最も小さかった 10^{-6} と 10^{-5} に決定した。次に、SLNN法による結果を用いて、DLNN法におけるIDWパラメータ m を決定した。现阶段では探索半径 $R=500$ mに固定し、 $m=1.0, 2.0, 4.0$ の3つの比較結果から、 $m=2.0$ とした。以上の検討結果を用いて、検証用データの推定を行った。なお、NNモデルの中間層数は9とし、更新計算にはADADELTA⁴⁾を用いた5,000万回の機械学習を実施している。

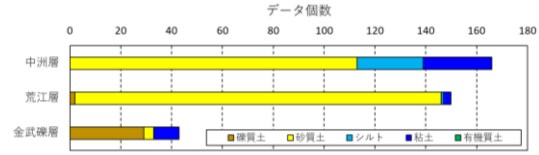
ONN法とDLNN法による検証用データの正解・不正解の分布と正解率の結果を表-1に示す。表には、L2正則化項の有無による結果の違いも示している。ONN法とDLNN法の結果より、学習用データの正解率はL2正則化の項を追加することによりやや低下するものの、検証用データの正解率は高くなっており、L2正則化の有効性が確認される。また、ONN法とDLNN法の検証用データの正解率は、DLNN法の方が10%程度上昇しており、技術者の作法を取り入れようとしたDLNN法による地層区分推定が優れている結果となっている。

BDを用いたNNによる地層区分推定の適用性について明らかにするために、福岡平野のある断面を対象とした検討を行った。全データ分け隔てなく学習時の判断材料とするONN法でも80%以上の正解率が得られたが、データを平均化して偏りを小さくし、推定位置に近いBDの情報ほど誤差が小さくなることを期待したDLNN法の方がより高い正解率となることが確かめられた。このことから、地盤情報の有する情報バイアスを考慮することが、より良い推定結果を得るためには必要であると思われる。

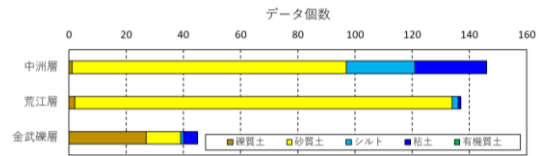
(3) 地層区分 AI 推定の精度評価

地盤情報を用いた地層区分 AI 推定では、正解率100%を得ることは困難である。これは地盤情報が有する情報バイアスによるところが大きい。機械学習の実施において、目指すべき正解率が明らかになれば、推定精度の目標値を定めることができ、推定結果の信頼性評価の上でも有効である。そこで、N値の分布にばらつきを有する2, 3層の模擬地盤を作成し、模擬地盤を用いた地層区分 AI 推定が正解率に与える影響を調査する。また、N値のばらつきを考慮した正解率の評価手法を提案し、その有効性を確認する。

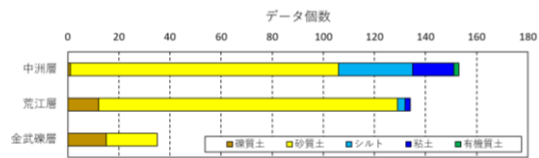
まず、機械学習における正解率の評価手法を次のように提案した。多層地盤でのN値頻度分布が図-6のような場合、任意のK層、および、全層の正解率は次式で与えられるとした。



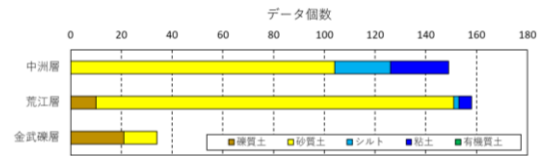
(a) Dataset A



(b) Dataset B



(c) Dataset C



(d) Dataset D

図-4 学習用データの層・土質毎のデータ個数

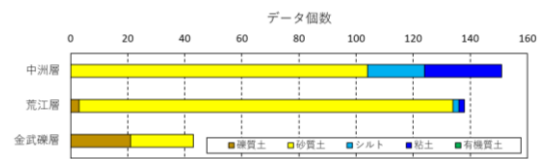
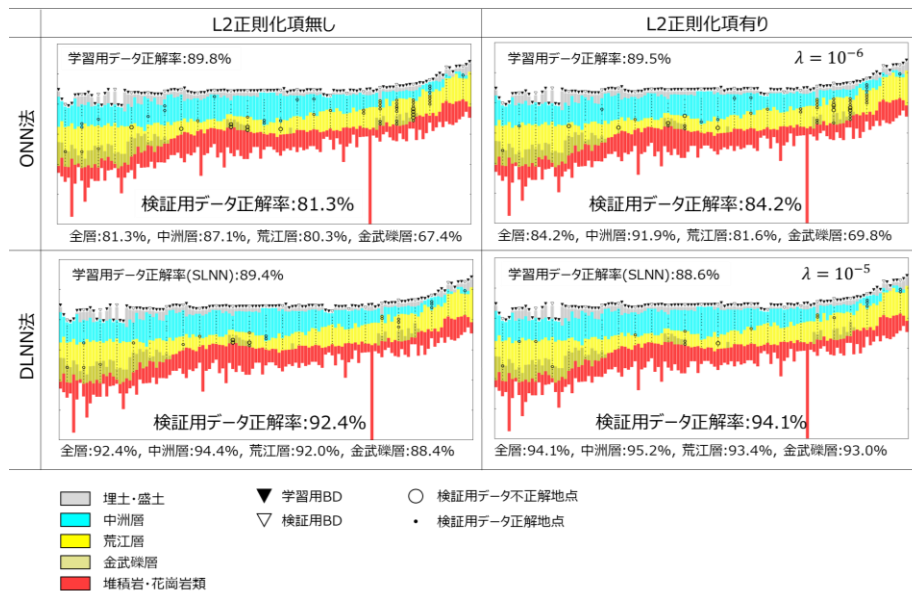


図-5 検証用データの層・土質毎のデータ個数 (Dataset X)

表-1 地層区分推定結果の比較



K 層の評価正解率: $\frac{N_{CK}}{N_K} \times 100$ ⑤

全層の評価正解率: $\frac{\sum_{K=1}^K N_{CK}}{\sum_{K=1}^K N_K} \times 100$ ⑥

ここで N_K は K 層の個数、 N_{CK} は機械学習の際に K 層になると予測する個数である。なお、N 値頻度分布において頻度分布曲線の交点がいく

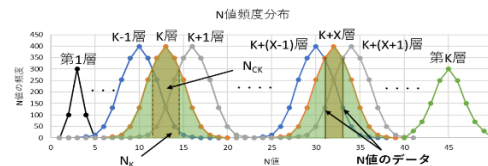


図-6 多層地盤での N 値頻度分布の模状図

つかの層で同数であった場合、同一の層数で除した個数を正解となる N 値個数としてカウントすることにする。この評価手法を用いて評価正解率を算出し、NN の機械学習により算出した NN 正解率を比較した結果を表-2～表-3 に示している。表-2 の評価正解率と NN 正解率を比較すると、同じ値になっていることがわかる。交点でデータが重複しない 3 層の NN 正解率を予測する評価手法の有効性を確かめることができ、4 層、5 層など、多層地盤にも適用可能であると考えられる。表-3 の評価正解率と NN 正解率を比較すると、正解率が高くなればなるほど誤差が小さくなることが分かった。交点でデータが重複する場合の NN 正解率を予測する評価手法は有効であると考えられる。

表-2 3 層の平均値を変える場合の評価正解率と NN 正解率による比較

	平均値			標準偏差			データ数(個)			評価正解率(%)			評価全体正解率(%)	NN正解率(%)			NN全体正解率(%)
	A層	B層	C層	A層	B層	C層	A層	B層	C層	A層	B層	C層	全層	A層	B層	C層	全層
CASE01-02A	10	11	12	2	2	2	2000	2000	2000	60	19.95	59.95	46.63	60	19.95	59.95	46.63
CASE02-02A	10	13	16	2	2	2	2000	2000	2000	77.4	54.75	77.35	69.83	77.4	54.75	77.35	69.83
CASE03-02A	10	15	20	2	2	2	2000	2000	2000	90.1	80.15	90.05	86.767	90.1	80.15	90.05	86.767

表-3 3 層の平均値を変える(交点でデータが重複する)場合の評価正解率と NN 正解率による比較

	平均値			標準偏差			データ数(個)			評価正解率(%)			評価全体正解率(%)	NN正解率(%)			NN全体正解率(%)
	A層	B層	C層	A層	B層	C層	A層	B層	C層	A層	B層	C層	全層	A層	B層	C層	全層
CASE01-12A	10	12	14	2	2	2	2000	2000	2000	68.7	37.35	68.65	58.23	60	37.35	77.35	58.23
CASE02-12A	10	14	18	2	2	2	2000	2000	2000	83.75	67.45	83.7	78.3	77.4	80.15	77.35	78.3
CASE03-12A	10	18	26	2	2	2	2000	2000	2000	97.95	95.8	97.85	97.2	96.7	98.3	96.6	97.2

交点でデータが重複する 3 層の模擬地盤の場合は、評価正解率と NN 正解率を比較すると正解率が高くなればなるほど、誤差が小さくなることが分かった。この結果は 2 層の場合も同様な結果となった。交点でデータが重複する 2, 3 層の模擬地盤の場合の NN 正解率を予測する評価手法は有効であると考えられる。

交点でデータが重複しない 3 層の模擬地盤の場合は、評価正解率と NN 正解率を比較すると正解率が同じ値であった。この結果は 2 層の場合も同様な結果となった。交点でデータが重複しない 2, 3 層の模擬地盤の場合の NN 正解率を予測する評価手法は有効性であることが確かめられた。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 村上 哲	4. 巻 70(5)
2. 論文標題 地盤情報DBに適用するNNによる地層区分推定法	5. 発行年 2022年
3. 雑誌名 地盤工学会誌	6. 最初と最後の頁 35-38
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計8件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 谷口大志, 村上 哲, 西 智美
2. 発表標題 ニューラルネットワークによる地層区分推定法適用基準の一般化
3. 学会等名 土木学会西部支部研究発表会
4. 発表年 2023年

1. 発表者名 湯屋泰成, 村上 哲, 西 智美
2. 発表標題 N値の空間分布推定手法の提案と福岡平野の天神・博多地区への適用
3. 学会等名 土木学会西部支部研究発表会
4. 発表年 2023年

1. 発表者名 村上哲
2. 発表標題 L2正則化を適用したNNによる各種地層区分推定法の比較
3. 学会等名 第57回地盤工学研究発表会講演
4. 発表年 2022年

1. 発表者名 金澤龍平, 村上 哲, 西 智美
2. 発表標題 NNによる地層区分推定法適用基準の構築
3. 学会等名 土木学会西部支部
4. 発表年 2022年

1. 発表者名 村上 哲
2. 発表標題 L2正則化を適用したD-Layer NN法による地層区分推定
3. 学会等名 土木学会西部支部
4. 発表年 2022年

1. 発表者名 遠藤 輝, 村上 哲, 西 智美
2. 発表標題 N値の空間分布推定手法の福岡平野沿岸域への適用性の検討
3. 学会等名 土木学会西部支部
4. 発表年 2022年

1. 発表者名 村上 哲, 赤間大地, 西 智美, 三輪 滋
2. 発表標題 NNによる地層区分推定のための NNモデルの構築手法
3. 学会等名 地盤工学会
4. 発表年 2021年

1. 発表者名 笠爽平, 村上哲, 西智美
2. 発表標題 NNを用いた地層区分推定法の適用基準確立のための模擬地盤を用いた検討
3. 学会等名 土木学会西部支部研究発表会講演概要集
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関