

通信用地下管路の震災時点検結果を基にした 機械学習による被害予測モデル検討

伊藤 陽¹・奥津 大²・古川 愛子³・庄司 学⁴・鈴木 崇伸⁵

¹正会員 日本電信電話株式会社 (〒 305-0051 茨城県つくば市花畑 1-7-1)

E-mail: akira.itou.dp@hco.ntt.co.jp

²正会員 日本電信電話 (同上)

E-mail: masaru.okutsu.ef@hco.ntt.co.jp

³正会員 京都大学准教授 工学研究科 都市社会工学専攻 (〒 615-8540 京都府京都市西京区京都大学桂)

E-mail: furukawa.aiko.3w@kyoto-u.ac.jp

⁴正会員 筑波大学教授 システム情報系 (〒 305-8573 茨城県つくば市天王台 1-1-1)

E-mail: gshoji@kz.tsukuba.ac.jp

⁵正会員 東洋大学教授 理工学部都市環境デザイン学科 (〒 350-8585 埼玉県川越市鯨井 2100)

E-mail: tsuzuki@toyo.jp

通信用管路の地震時の被害を事前に予測することが、震災対策において重要となる。本稿では地震時の管路被害予測に機械学習により構築される予測モデルが有効であるか検証した。1995 年兵庫県南部地震, 2004 年新潟県中越地震, 2007 年新潟県中越沖地震, 2011 年東北地方太平洋沖地震の 4 種の地震における管路点検結果を学習用データとしてモデルを構築した。このモデルは 4 地震のテストデータに対しては十分な分類性能を持っていたが、熊本地震の被害を予測を行うためにはさらなる改善が必要であると考えられる。これに対しモデル構築の際の学習用データを直下型と海溝型に分類してのモデル構築を試みたが、双方ともに大きな改善は見られず使用データの精査など追加の検討が必要であるといえる。

Key Words: *Seismic damage prediction, Telecommunication conduit, Machine learning, Gradient boosting decision tree*

1. はじめに

地下に敷設された通信ケーブルを防護する通信用地下管路は総延長 62 万 km におよび、その多くは 1960 年代から 70 年代に建設されたものであり伸縮性能のない継手を有する管路も多数現存している。通信においては管路は可能な限り更改せず利用し続ける事を目標としており、震災対策も建設年に関わらず真に地震時に被害を受けるものに対してのみ行う必要がある¹⁾。すなわち、地震時の管路被害を予測する技術が重要となる。

地下管路の地震被害を予測する研究は多く行われており、例えば水道管では磯山らが兵庫県南部地震における芦屋市と西宮市の水道管被害データより標準被害率曲線を作成し²⁾、これに丸山・山崎は 2004 年新潟県中越沖地震, 2007 年能登半島地震, 2007 年中越沖地震の管路データを加え被害予測式を立案している³⁾。さらに鉏田・大野は、東北地方太平洋沖地震における宮城県北西部の管路被害から管種・口径の補正係数を提案している⁴⁾。また通信用管路の被害予測においては庄司らが管種、亘長、微地形区分で分類するスクリーニング手法および地震被害関数を構築している⁵⁾。一方で被害予測に機械学習を用いる研究も近年見られ、滝らは被害推定を異常検知の問題としてとらえ教師無学習の

一つである密度比推定を用いて 50 m メッシュでの異常度を推定している⁶⁾。また能島・大西は複数アルゴリズムでの予測の基礎検討を行っている⁷⁾。

本稿では、通信用地下管路の地震時被害予測に機械学習による予測モデル作成が適用可能であるかを確認した。モデル作成においては決定木による勾配ブースティングの一つである XGboost⁸⁾ を利用し、学習用およびテストのデータは、1995 年兵庫県南部地震, 2004 年新潟県中越地震, 2007 年新潟県中越沖地震, 2011 年東北地方太平洋沖地震, 2016 年熊本地震における通信用地下管路の点検に基づく被害有無データを用いた。まず全データを用いて学習およびテストを行い、テストの結果および変数寄与度を確認することでモデルの有効性を把握する。その上で同様の手法で 1995 年兵庫県南部地震から 2011 年東北地方太平洋沖地震までのデータを基にモデルを構築し、熊本地震のデータをテストデータとして突合し、被害が予測可能であるかを検証した。

2. 機械学習によるモデル構築

(1) 基となるデータベース

データベースは、地震における点検結果と当該箇所の地震動情報、設備情報、地盤情報から構成される。対

表-1 利用した変数

項目	概要
管種	鋼管 (ねじ継手) / 鋼管 (差込継手) ⁹⁾ ビニル管 (接着継手) / ビニル管 (差込継手) ¹⁰⁾ 鋳鉄管 (ねじ継手)
亘長	当該管路のスパン長
経過年数	建設から地震までの年数
AVS30	地下 30m までの平均 S 波速度, J-SHIS ¹¹⁾ の微地形区分に基づく値にボーリングデータ ¹²⁾ を統合したもの
基本固有周期	地盤の固有周期, J-SHIS の微地形区分に基づく値 ¹³⁾ にボーリングデータを統合したもの
推定人工平坦化地	盛土等, 人工的に平坦化された ¹⁴⁾ と推定される箇所
平均標高	基盤地図情報数値標高モデル 10m メッシュより作成された 250m メッシュの平均標高 ¹⁵⁾
平均傾斜角	基盤地図情報数値標高モデル 10m メッシュより作成された 250m メッシュの平均傾斜角
PGV	K-NET, KIK-net ¹⁶⁾ で観測された最大速度を Kriging 法で 250m メッシュにて推定
PGA	K-NET, KIK-net で観測された最大加速度を Kriging 法で 250m メッシュにて推定
計測震度	K-NET, KIK-net で観測された計測震度を Kriging 法で 250m メッシュにて推定
SI 値	K-NET, KIK-net で観測されたスペクトル強度を Kriging 法で 250m メッシュにて推定
換算変位	推定された PGV^2/PGA で変換される最大変位に相当する値.
等価卓越周期	推定された $2\pi PGV/PGA$ で変換される周期
地盤ひずみ	高圧ガス導管耐震設計指針に基づき算出されるひずみ

象とする地震は前節で示した通り, 1995 年兵庫県南部地震, 2004 年新潟県中越地震, 2007 年新潟県中越沖地震, 2011 年東北地方太平洋沖地震, 2016 年熊本地震の 5 地震である. 点検は規定のサイズの治具による通過試験で通過しなかった管路に対してパイプカメラを挿入し内部状況を確認することで行う. 主な被害ケースとしては, 継手の離脱及びこれに伴う土砂流入, 継手の屈曲, 管体の扁平がある. 仮に通過しなかった箇所で先述のような様子が確認できず, 錆や工具等によると考えられる刺突の痕跡により治具が不通過となっている場合, 当該箇所は被害なしとして集計する. 点検が実施された範囲は, 震度階で 6 弱以上が観測された自治体を含む, ケーブル異常や路面の変状等が確認されている収容局の設備であり, 全数の点検データではない. 管路はマンホールとマンホールを繋ぐ形で建設されており, 複数条存在するケースが多いが, 主に予備として用意されたケーブルの入っていない管路を対象に点検された結果を利用する. 各スパンに紐づく説明変数は表-1 に示す 15 種類のものである. 本検討では, 設備に関するデータ (管路種別など), 地盤に関するデータ (AVS30 など), 地震動に関するデータ (計測震度など) の三種について, それぞれ全地震の点検データにおいて取得可能であった変数を利用する. また各地震におけるデータ数を表-2 に示す.

(2) 機械学習手法

今回利用した手法は決定木を用いた勾配ブースティング (GBDT) を利用する. ブースティングは分類のために設計されたアルゴリズムであり, 弱分類器を大量に出力し, これを組み合わせで合議し予測結果を出力する手法であるが, あるステップ k での弱分類器の生

表-2 各地震のデータ数

地震名	スパン数	内被害有スパン数
兵庫県南部地震	3955	861
新潟県中越地震	2130	55
新潟県中越沖地震	761	102
東北地方太平洋沖地震	19211	162
熊本地震	6977	217

成において, 前ステップ $k-1$ までの分類器が正しく分類できなかった学習データに重点を置いて学習することが特徴である¹⁷⁾. ここで, m 個の変数を持つ n 個のレコードよりなるデータセットを想定する. このとき i ($i = 1, 2, \dots, n$) 番目のレコードの変数が m 次元のベクトル $\mathbf{x}_i$ と表され, i 番目のレコードの被害有無を 0, 1 の二値で y_i と表されるとする. 今回の場合, 15 の変数を用いるため $m = 15$ となる. ここで決定木 f を考えると, $\mathbf{x}_i$ が入力されたとき, 木の J 個の重複の無い頂点 R_j に割り当てられた定数 w_j ($j = 1, 2, \dots, J$) が出力されるモデルであり,

$$f(\mathbf{x}_i) = \sum_{j=1}^J w_j I(\mathbf{x}_i \in R_j) \quad (1)$$

と表現することができる. この時, I は指示関数であり, 条件に該当する際には 1 をそれ以外には 0 を返す関数である. このため K 個の決定木 f の合議 ϕ によって算出される予測値 $\hat{y}_i$ は,

$$\hat{y}_i = \phi(\mathbf{x}_i) = \sum_{k=1}^K f_k(\mathbf{x}_i) \quad (2)$$

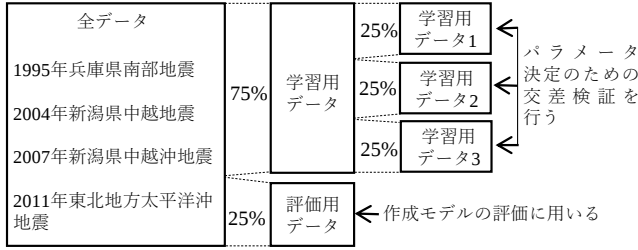


図-1 モデル構築におけるデータ分割

となる．終端頂点数 J の最大値，決定木の終端頂点で得られる値 w_j の最小値，合議に用いる決定木の合計数 K は，モデルによる予測値である $\hat{y}_i$ に影響するパラメータとなる．ただし学習の際には，過学習防止のために得られた木の出力結果を 0 から 1 の間の値をとる学習率 η で縮小する．すなわち

$$\hat{y}_i^k = \hat{y}_i^{k-1} + \eta f_k(\mathbf{x}_i) \quad (3)$$

となる．

この時，目的関数 $\mathcal{L}(\phi)$ を以下のように定義する．

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (4)$$

ここで， $l(\hat{y}_i, y_i)$ は $\hat{y}_i$ と y_i の間の差を測る微分可能な損失関数であり，頂点の数 J と頂点で得られる値のベクトル $\mathbf{w} = (w_1, w_2, \dots, w_J)$ を有する決定木 f に対して $\Omega(f)$ は

$$\Omega(f) = \gamma J + \frac{1}{2} \lambda \|\mathbf{w}\|^2 \quad (5)$$

と表現される罰則項である． γ は木の大きさに対する罰則を与える係数で， λ は木の各頂点が返す値の大きさに対する罰則を与える係数である．木が大きくなり，返す値が小さくなると細かく表現することが可能となるが，同時に過学習を起こしやすくなるため，これらもパラメータとして最適化する必要がある．

(3) モデル構築

今回，熊本地震データの予測の前に構築したモデルの評価を行うために兵庫県南部地震，新潟県中越地震，新潟県中越沖地震，東北地方太平洋沖地震の4地震のデータの内 20% を評価用のテストデータとして分割した．すなわち，モデル構築のための学習は残りの 75% のデータを用い，モデルの評価は 25% のテストデータを用いて行った．学習の際には 75% のデータを3分割し交差検証して作ったモデルの受信者動作特性曲線 (ROC 曲線) 下の面積の値である Area Under the Curve (AUC) の値が最大化するようにパラメータを決める．ROC 曲線の横軸は実際には被害無のデータを被害有と判定し

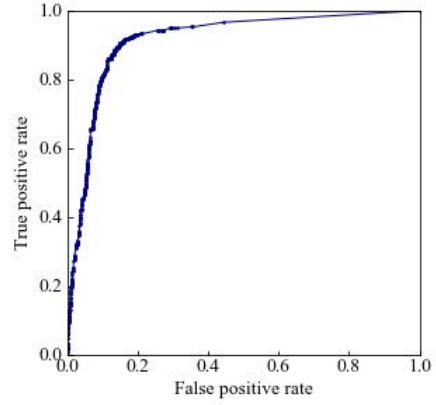


図-2 作成モデルテスト結果 (ROC 曲線)

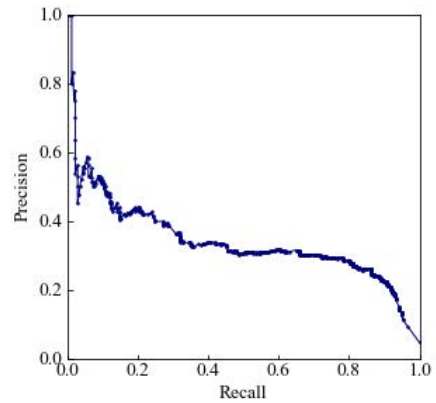


図-3 作成モデルテスト結果 (PR 曲線)

てしまった割合である偽陽性率であり，縦軸は実際に被害有のデータを正しく被害有と判定できた割合である真陽性率である．ROC 曲線は被害有無を分ける閾値を変化させることでプロットされ，ROC-AUC は 0 から 1 の値を取り 1 に近いほど良好なモデルであることを示す．モデル構築の際のデータの利用について図-1 に示す．各分割の際には，全データに占める被害有データが等しい割合になるように分割しているが，各地震の割合はランダムとなっている．

GBDT において先述の γ ， λ ， J の最大値 J_{max} ， w_j の最小値 w_{jmin} ， K ，および学習率 η を定める必要があるが，今回は $\eta = 0.001$ ，と設定し， γ ， λ ， J_{max} ， w_{jmin} は調整はベイズ最適化アルゴリズムの一種である Tree-structured Parzen Estimator を基にしている Optuna¹⁸⁾ を利用した．

3. 結果と考察

(1) テストデータによるモデル評価

兵庫県南部地震，新潟県中越地震，新潟県中越沖地震，東北地方太平洋沖地震の4地震の全データの内 75%

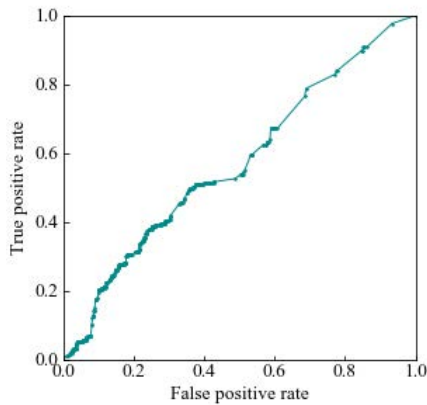


図-4 作成モデルによる熊本地震被害予測結果 (ROC 曲線)

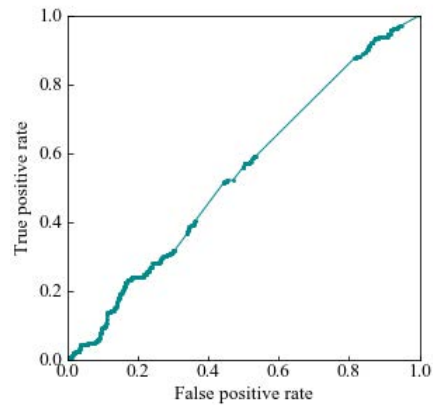


図-5 直下モデルによる熊本地震被害予測結果 (ROC 曲線)

を用いて構築したモデルに対し、残りのデータを適用してテストした。ROC 曲線を図-2 に示す。また同時に真陽性率を横軸に、被害有と予測したデータの内実際に被害有であるデータの割合を示す適合率を縦軸にとった Precision-Recall 曲線 (PR 曲線) を図-3 に示す。PR 曲線の AUC である PR-AUC も 1 に近いほど良好なモデルであることを示すが、PR-AUC は被害有、被害無レコード数に偏りのある場合の評価指標として有効である。それぞれの AUC は、ROC-AUC が 0.92, PR-AUC が 0.34 であった。ここで、被害有データに対する被害無データの割合である skew ratio は 21.1 であり偏りが大きいといえる。Jeni ら¹⁹⁾ は skew ratio が異なるデータに対し、異なる相対的な誤分類率を有する二値分類モデルによる評価を行っている。これに基づき skew ratio が 21.1 の際に各 AUC の値が今回の値である場合、モデルの精度は約 10% の誤分類率を有するものとなった。

(2) 4 地震モデルでの熊本予測

先述の 4 地震を基にしたモデルを、熊本地震において点検された管路のデータに対して適用した結果の ROC 曲線を図-4 に示す。予測の際には、4 地震のモデルの平均および標準偏差で熊本地震被害データを標準化している。ROC-AUC は 0.57 であり、(1) 項の結果である 0.92 と比較すると被害有無を正確に分類することができていない。これよりモデル構築に用いたデータと同一の地震内であれば十分に分類できている一方で、モデルにとって新規の地震である熊本地震が分類できていないことが判る。

熊本地震の被害を予測できていない要因の一つとして、直下型地震と海溝型地震のデータを混合させてモデルを構築していることが影響している可能性があるため、これを検証する。モデルを兵庫県南部地震・新潟県中越地震・新潟県中越沖地震の三種より作成した直下型モデルと東北地方太平洋沖地震を基に作成した海溝

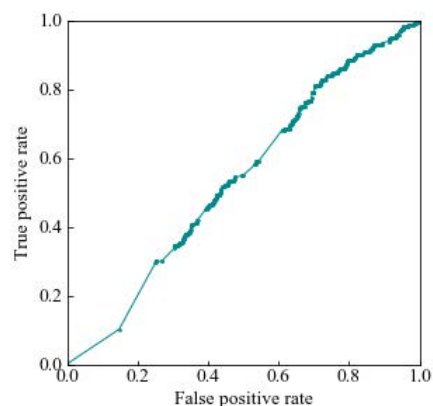


図-6 海溝モデルによる熊本地震被害予測結果 (ROC 曲線)

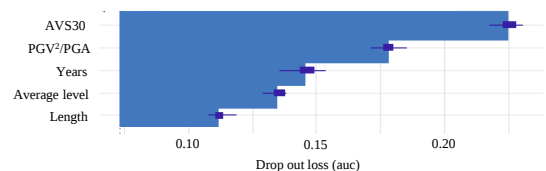


図-7 直下型モデルにおける変数寄与度 (AUCPFI)

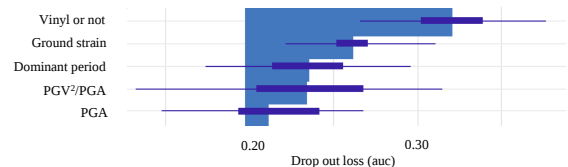


図-8 海溝型モデルにおける変数寄与度 (AUCPFI)

型モデルの二種類で熊本地震の被害を予測した。直下型モデル結果を図-5 に、海溝型モデルの精度は図-6 に示す。直下型モデル、海溝型モデルともに ROC-AUC は 0.54 であり、4 地震を基にしたモデルより予測精度は下がっていると言える。これらのモデルがどのような

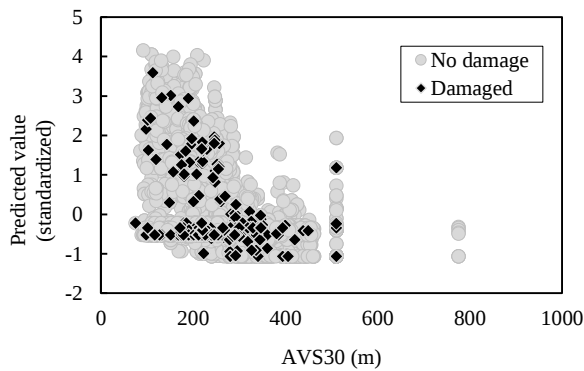


図-9 AVS30 に対する直下型モデルの予測値

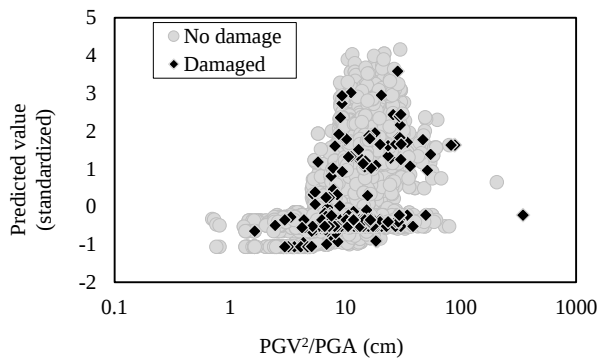


図-10 換算変位に対する直下型モデルの予測値

特徴を有しているか把握するため、Permutation feature importance を確認した²⁰⁾。Permutation feature importance は一つの変数をランダムに入れ替え、モデルの性能がどの程度低下を確認する手法であり、入れ替えた際に性能が大きく下がる変数はモデルに大きく寄与している変数であると言える。結果を図-7 および図-8 に示す。直下型モデルにおいては AVS30 や換算変位などが、海溝型モデルにおいては接着式ビニル管であるか否かが大きく影響していると言える。

この結果から、AVS30 に対する直下型モデルに熊本地震のデータを適用させた結果の予測値および被害有無をプロットしたものを図-9 に示す。予測値は AVS30 が 200m/s 付近までは相対的に高い値を示す傾向にあるが、熊本地震においては 300~500m/s にも被害有が多く存在している。これは、熊本ではローム台地や阿蘇山付近の山麓地が多く存在しており、これらの箇所でも被害を受けていることから既往の地震と異なる傾向にあり、予測精度が低下しているといえる。また、換算変位についても同様にプロットした図を図-10 に示す。ここから 8~30cm において予測値が高くなっていることが判るが、熊本地震においては断層変位による被害

も見られることから、換算変位による予測では不足する箇所があることも考えられる。

また、海洋型モデルについては旧式のビニル管であるか否かが支配的であり、ビニル管である場合の予測値の平均が 1.45 に対し、ビニル管でない場合の予測値の平均は -0.64 であった。東北地方太平洋沖地震において点検された管種に偏りがあったために、被害予測に資するモデルになっていない可能性がある。

4. まとめ

本稿では、過去の地震時の点検データから機械学習を用いたモデルを作成し、学習させていない熊本地震のデータに対して正しく予測可能か検証した。

1. 勾配ブースティングを用いて、兵庫県南部地震、新潟県中越地震、新潟県中越沖地震、東北地方太平洋沖地震の4地震を基にしたモデルを作成しテストすると、誤分類率 10%程度の予測精度を持つモデルを構築することができた。
2. 4地震を基にしたモデルを用いて熊本地震の被害を予測したが、ROC-AUC が 0.57 と十分に予測できているモデルとはいえないモデルとなった
3. 直下型モデルと海溝型モデルを用いて熊本地震の被害を予測したが、ROC-AUC 0.54 と予測できていないモデルとなった。内容を分析すると、熊本地震の微地形区分がこれまでと違う傾向であったこと、断層変位による被害がパラメータに入っていないことなどが原因として挙げられる。また海溝型モデルは点検データの偏りが影響していると考えられる。

今後、データを被害原因や管路種別ごとに分類しモデルを構築する。また、変数の増減などを行うことで予測精度が向上するか検討する。

参考文献

- 1) 杉野文秀, 政倉浩志: 通信基盤設備の安心・安全かつ経済的運用を実現する維持管理技術, NTT 技術ジャーナル, Vol.8, pp.8-11, 2014.
- 2) 磯山龍二, 石田栄介, 湯根清二, 白水暢: 水道管路の地震被害予測に関する研究, 水道協会雑誌, Vol. 67, No. 2, pp.25-40, 1998.
- 3) 丸山喜久, 山崎文雄: 近年の地震被害データを加味したマクロな配水管被害予測式の改良, 土木学会論文集 A1, Vol. 65, No. 1, pp.565-574, 2009.
- 4) 鯉田泰子, 大野顕大: 東北地方太平洋沖地震における宮城県北西部の管路被害の特徴, 土木学会論文集 A1, Vol. 69, No. 4, I_734-I_741, 2013.
- 5) 庄司学, 宮崎史倫, 若竹雅人, 伊藤陽, 鈴木崇伸: 通信埋設管路の地震対策に活用するスクリーニング手法の提案及び地震被害関数の構築, 土木学会論文集 A1, Vol.72, No.4, I_523-I_541, 2016.
- 6) 滝勇太, 猪俣渉, 丸山喜久: 機械学習技術を活用した低圧

- ガス導管被害推定手法の検討, 第9回インフラ・ライフライン減災対策シンポジウム講演集, pp.71-75, 2019.
- 7) 能島暢呂, 大西克茂: 地中埋設管の地震被害予測への機械学習の適用に関する基礎的検討, 東濃地震科学研究所報告, Vol. 43, pp.23-31, 2019.
 - 8) Tianqi Chen, Carlos Guestrin: XGBoost: A Scalable Tree Boosting System, arXiv: 1603.02754
 - 9) JFE スチール株式会社: 差込み継手塗覆装鋼管, <https://www.jfe-steel.co.jp/products/koukan/catalog/e1j-024.pdf>, (2020.9.4 閲覧)
 - 10) クボタケミックス: https://www.kubota-chemix.co.jp/products/pvc_pipes/for_communication_cable/item_1555 (2020.9.4 閲覧)
 - 11) 防災科学技術研究所: <http://www.j-shis.bosai.go.jp/> (2020.9.4 閲覧)
 - 12) 末富岩雄, 石田栄介, 福島康宏, 磯山龍二, 澤田純男: 地形分類とボーリングデータの統合処理による地盤増幅度評価と2004年新潟県中越地震における地震動分布の推定, 日本地震工学論文集, Vol.7, No.3, pp.1-12, 2007.
 - 13) 先名重樹, 翠川三郎: 地形・地盤分類に基づく地震動のスペクトル増幅率の推定, 日本地震工学論文集, Vol.9, No.4, pp.11-25, 2009.
 - 14) A. Ito, M. Okutsu, T. Suzuki, G. Shoji: Development of Estimation Method of Seismic Damage for Telecommunication Conduit, 7th ICEGE proceedings, pp.3062-3069, 2019.
 - 15) 国土交通省 国土数値情報ダウンロードサービス: <http://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-G04-d.html> (2020.9.4 閲覧)
 - 16) 防災科学技術研究所 強震観測網: <http://www.kyoshin.bosai.go.jp/kyoshin/> (2020.9.4 閲覧)
 - 17) Trevor Hastie, Robert Tibshirani, Jerome Friedman: The Elements of Statistical Learning: Data Mining Inference and Prediction 2014, 杉山将, 井出剛, 神島敏弘, 栗田多喜男, 前田英作 監訳, 統計的学習の基礎 - データマイニング・推論・予測 -, 共立出版株式会社, 2014.
 - 18) Preferred Networks: <https://preferred.jp/ja/projects/optuna/> (2020.8.31. 閲覧)
 - 19) László A. Jeni, Jeffrey F. Cohn, Fernando De La Torre: Facing imbalanced Data Recommendations for the Use of Performance Metrics, Conference on Affective Computing and Intelligent Interaction, pp.245-251, 2013.
 - 20) Andreé Altmann, Laura Toloşi, Oliver Sander, and Thomas Lengauer: Permutation importance: a corrected feature importance measure, Bioinformatics, Vol. 26, Issue10, pp.1340-1347, 2010.

DAMAGE PREDICTION MODEL BY MACHINE LEARNING BASED ON THE RESULTS OF INSPECTION OF CONDUITS FOR TELECOMMUNICATION

Akira ITO, Masaru OKUTSU, Aiko FURUKAWA, Gaku SHOJI, Takanobu SUZUKI

It is important for earthquake countermeasures to predict the damage of communication pipelines in advance. This paper examines the effectiveness of a prediction model built by machine learning for predicting damage to pipelines during earthquakes. It was built. Although the model had good classification performance for the four earthquakes tested, it could hardly predict the damage of the Kumamoto earthquake. In addition, we classified the training data into two types, nadir and trench, and found that both types of models had low classification performance, but the trench model was particularly unpredictable.