

AI を用いた異常予兆検知システムの開発

(2) 2段階オートエンコーダの発電所データでの評価

Development of Anomaly Detection System using AI
(2) Evaluation of two-stage autoencoder using power plant data

株式会社 東芝	田口 安則	Yasunori TAGUCHI	Member
株式会社 東芝	内藤 晋	Susumu NAITO	Member
株式会社 東芝	加藤 佑一	Yuichi KATO	Non-member
株式会社 東芝	中田 康太	Kouta NAKATA	Non-member
東芝エネルギーシ ステムズ株式会社	富永 真哉	Shinya TOMINAGA	Non-member
東芝エネルギーシ ステムズ株式会社	名倉 伊作	Isaku NAGURA	Non-member
東芝エネルギーシ ステムズ株式会社	三宅 亮太	Ryota MIYAKE	Non-member
東芝エネルギーシ ステムズ株式会社	青木 俊夫	Toshio AOKI	Non-member
東芝エネルギーシ ステムズ株式会社	寺門 優介	Yusuke TERAKADO	Non-member
東芝エネルギーシ ステムズ株式会社	高戸 直之	Naoyuki TAKADO	Non-member
東芝エネルギーシ ステムズ株式会社	高森 由己夫	Yukio TAKAMORI	Non-member
東芝デジタル&コ ンサルティング株 式会社	大熊 栄一	Eiichi OOKUMA	Non-member

In power plants, a large number of sensors are installed and the measured values are used to control and operate the plants. Although operators monitor the measured values for daily operation and maintenance, the number that can be visually confirmed is limited. So far, we have proposed two-stage autoencoder for anomaly detection. To evaluate its effectiveness, it and simple autoencoder as a comparison method were applied to the operational data of a power plant. As a result, deterioration of a thermometer in the boiler combustion chamber was detected by both methods. The false positive ratio of the two-stage autoencoder was about 70% of that of the simple autoencoder. These results showed practicality of the two-stage autoencoder.

Keywords: anomaly detection, two-stage autoencoder

1. はじめに

連絡先: 田口安則、〒212-8582 神奈川県川崎市幸区小
向東芝町1番地、(株)東芝 研究開発センター、
E-mail: yasunori.taguchi@toshiba.co.jp

社会に電力を供給する発電プラントにおいては、常
時、大量のプロセスデータが収集、監視され、プラント
が運転、制御されている。プラント内の各部分は連動する
ため、各プロセス量の間には相関関係がある(図1)。
発電プラントに代表されるインフラ設備には、高い信

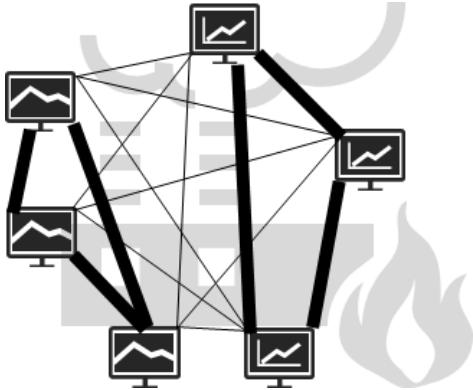


図1 発電プラント内のセンサとセンサ間の相関関係

頼性と、安定的な運転が求められる。そのため、異常またはその予兆を、できるだけ早く検知できることが望ましい。

運転員は、各プロセス値の上下限値の超過や、他の関係するプロセス値との相関関係の崩れを監視する。しかし、限られた人数で同時に監視できる数には限りがある。そのため、多数のプロセス値を同時に自動で監視するシステムが求められる。

近年、機械学習に基づく異常予兆検知が数多く研究されている⁴⁾。中でも、最も基本的で、かつ、検知性能が高い方式として知られているのがオートエンコーダ（以下、AE）⁵⁾に基づく方式である。この方式をプラントの異常予兆検知に応用する場合、例えば、監視対象とするシステムにおける複数のプロセスの瞬間値を並べたベクトルをAEに入力する。AEが再構成した出力ベクトルの入力ベクトルに対する再構成誤差は、互いに関係するプロセス値間での相関関係が崩れたときに大きくなりがちなため、当該システムの異常予兆を検知できる。しかし、プロセス値の時間方向の相関関係が崩れる異常予兆を検知できない。

それに対して、我々はこれまでに、時間方向の情報も活用する2段階AEとそれに基づく異常予兆検知システムを提案した⁶⁻⁸⁾。今回、この方式の有効性を確認するために、稼働中の発電所のプロセスデータで評価した。その結果、ボイラ燃焼室に設置されたある温度計の劣化の検知に成功した。この劣化は、関係するプロセス値との相関関係を崩すため、単純なAEでも検知されたものの、誤検知率が2段階AEの方が低かった。これにより、2段階AEの実用性を確認した。

以降、2、3で、単純なAEと2段階AEに基づく異常予兆検知について、それぞれ説明する。4で、両AEに基づく方式の発電所データでの評価結果を示し、5

で、まとめる。

2. 単純なAEに基づく異常予兆検知

2.1 単純なAE

AE⁹⁾は、入力ベクトル \mathbf{x} から出力ベクトル \mathbf{y} を生成するニューラルネットワークの一種である。異常予兆検知に応用されたり、特徴ベクトルの次元圧縮に用いられたりすることが多い。

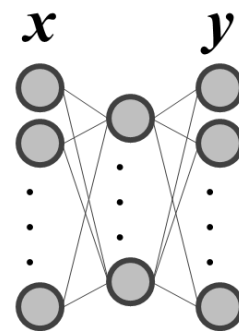
図2に、最も単純なAE（以下、単純AE）のネットワーク構造を示す。ここで、丸印がノードを表し、ノード間を結ぶ線がリンクを表す。単純AEは、入力層と中間層、出力層の3層からなる。中間層のノード数は入力層よりも少なく、出力層のノード数は入力層と同じである。ノードはバイアスを持ち、リンクは重みを持つ。

重みやバイアスは、AEを活用する前に学習する。この事前学習として、訓練データ $\{\mathbf{x}_n\}_{n=1,2,\dots,N}$ に対する再構成誤差の総和 $\sum_{n=1}^N \|\mathbf{y}_n - \mathbf{x}_n\|^2$ を最小化するものを求める。

中間層のノード数が入力層よりも少ないため、入力ベクトルの情報が中間層で圧縮される。その圧縮データから、出力ベクトルが再構成される。

データを圧縮する部分、すなわち、入力層と中間層からなる部分は、エンコーダと呼ばれる。エンコーダによる圧縮データから出力ベクトルを再構成する部分、すなわち、中間層と出力層からなる部分は、デコーダと呼ばれる。

学習済みAEの出力ベクトル \mathbf{y} は、訓練データにおける出現頻度が高い入力ベクトル \mathbf{x} に対する再構成誤差 $\|\mathbf{y} - \mathbf{x}\|^2$ が小さくなる。逆に、訓練データにおける出現頻度が低い \mathbf{x} に対しては、再構成誤差が小さくなる保証がない。



入力層 中間層 出力層

図2 最も単純なAEのネットワーク構造

2.2 単純なAEに基づく異常予兆検知

この性質を利用することで、AEで異常予兆を検知できる。AEに基づく異常予兆検知は、図3に示す事前学習ステップと、図4に示す異常予兆検知ステップからなる。プラントのあるシステムを監視する場合を例にとり、各ステップの処理を説明する。

事前学習ステップ（図3）は、サブステップ(T1)～(T3)からなる。以下に各サブステップでの処理を示す。

(T1) AEのネットワーク構造を図2に定める。入力ベクトル \mathbf{x} は、監視対象システムの状態を表す複数のプロセスの瞬間値を並べたベクトルとする。

(T2) 訓練データとして、プラントが正常な場合のデータを準備する。

(T3) AEの重みやバイアスを学習する。

異常予兆検知ステップ（図4）は、サブステップ(D1)～(D4)からなる。以下に各サブステップでの処理を示す。

(D1) 監視する瞬間の入力ベクトル \mathbf{x} を取得する。

(D2) \mathbf{x} を学習済みAEに入力し、出力ベクトル \mathbf{y} を取得する。

(D3) 再構成誤差 $\|\mathbf{y} - \mathbf{x}\|^2$ が所定の閾値より大きいか否かを判定する。大きかった場合、監視対象システムのその瞬間の状態が異常またはその予兆であるとみなす。異常予兆があるプロセス値を特定するには、各プロセス値の再構成値からの偏差の絶対値 $|y_i - x_i|$ が所定の閾値を超過したプロセス値を特定する。ここで、 x_i, y_i は \mathbf{x}, \mathbf{y} の i 番目の要素を表す。閾値を超過したプロセス値のトレンドグラフの例を図5示す。ここで、横軸が時間を表し、縦軸が値を表す。この例では、図中の途中で偏差が大きくなり、異常予兆として検知される。

(D4) 監視の継続を決定する。継続する場合は(D1)に戻り、継続しない場合は異常予兆検知ステップを終了する。

これにより、互いに関係するプロセス値間の相関関係の崩れが異常予兆として検知される。しかし、プロセス値の時間方向の相関関係が崩れる異常予兆を検知できない。

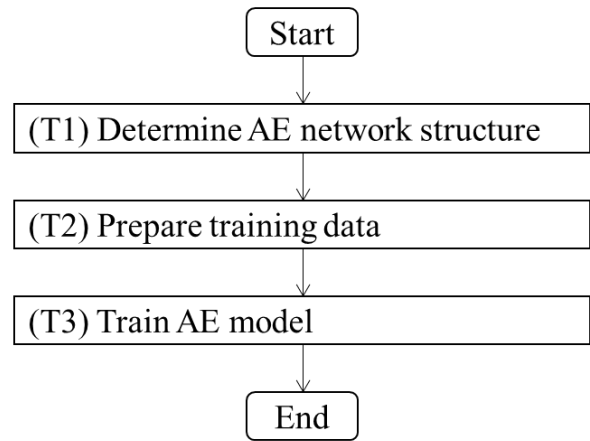


図3 事前学習ステップ

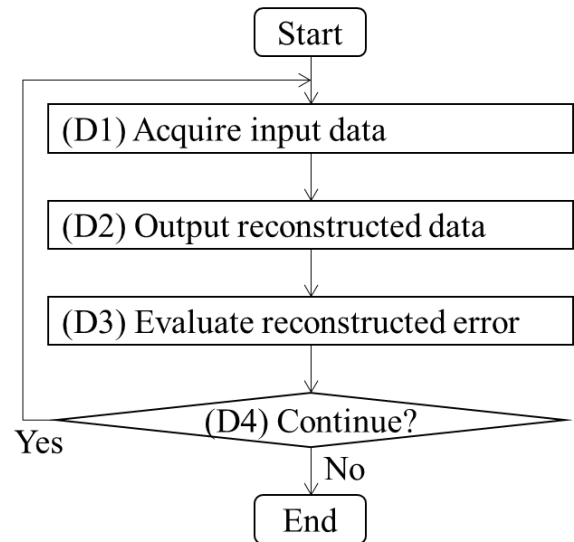


図4 異常予兆検知ステップ

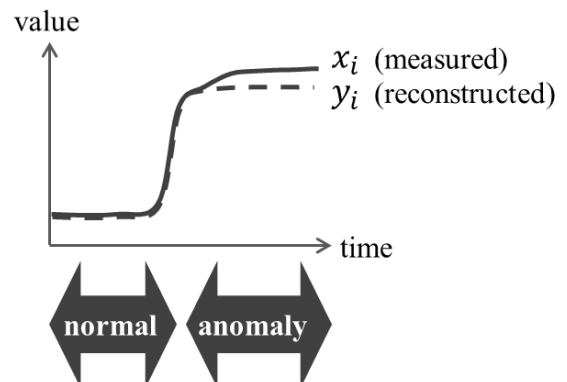


図5 トрендグラフの例

3. 2段階 AE に基づく異常予兆検知

3.1 2段階 AE

2段階 AE³⁾は、時系列データを扱うニューラルネットワークであり、内部に2つの AE を有する。図6に、ネットワーク構造を示す。入力ベクトル x は注目時刻を含む複数の時刻の値を要素に持つ時間窓データであり、出力ベクトル y は注目時刻の瞬間値である。2つある AE のうち、図中左側の1段階目 AE は時間窓データを扱い、図中右側の2段階目 AE は瞬間値を扱う。

2段階 AE の重みやバイアスは、訓練データを用いて2段階で学習する。まず、1段階目 AE のみを訓練する。そして、1段階目 AE の重みとバイアスを固定し、2段階目 AE を訓練する。

図7に、2段階 AE 内の各 AE が再構成する信号成分のイメージ図を示す。学習済みの1段階目 AE は、複数のプロセスの複数時刻の値を並べたベクトルに対する再構成誤差が最小化されたものであるため、支配的な誤差となる大まかな変動を再構成する。一方、学習済みの2段階目 AE は、注目時刻の偏差ベクトルに対する再構成誤差が最小化されたものであるため、細かい変動を再構成する。

2段階 AE を発電プラントデータに適用した場合、1段階目 AE では、プラントの運転状態に応じて変化する、各部プロセス量の時間的変動が主に再構成される(図8)。一方、2段階目 AE では、動的機器のローカルな微小振動による、プロセス量の微小変動が主に再構成される(図9)。

学習済み2段階 AE の出力ベクトル y は、訓練データにおける出現頻度が高い入力ベクトル x に対する再構成誤差 $\|y - g(x)\|^2$ が小さくなる。ここで、 g は注目時刻の値を抽出して並べたベクトルを出力する関数である。逆に、訓練データにおける出現頻度が低い x に対しては、再構成誤差が小さくなる保証がない。

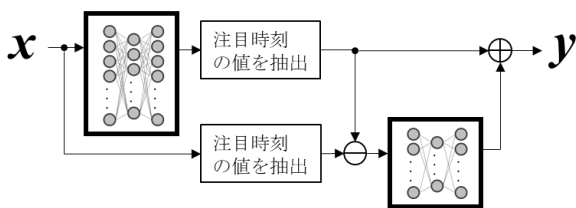


図6 2段階 AE のネットワーク構造

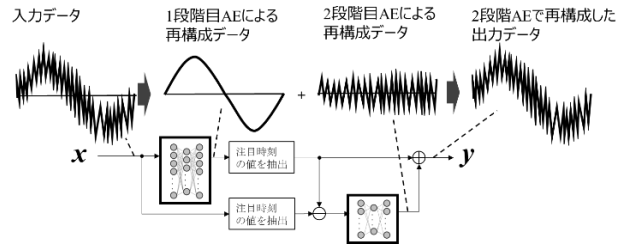


図7 2段階 AE 内の各 AE が再構成する信号成分のイメージ図

プラント運転状態に応じて変化する、各部プロセス量の時間的変動特性（動特性）



図8 プラントデータに対し1段階目 AE が再構成する信号成分

動的機器のローカルな微小振動による、プロセス量の微小変動特性

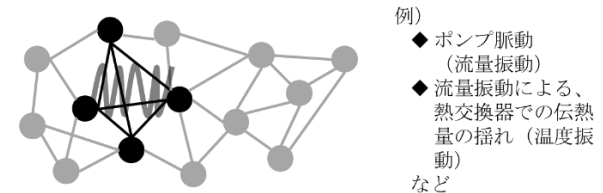


図9 プラントデータに対し2段階目 AE が再構成する信号成分

3.1 2段階 AE に基づく異常予兆検知

この性質を利用することで、2段階 AE に基づいて異常予兆を検知する。2段階 AE に基づく異常予兆検知も2.2の単純 AE に基づく異常予兆検知と同様に、図3に示した事前学習ステップと、図4に示した異常予兆検知ステップからなる。異なるのは、ネットワーク構造と入力ベクトル x 、および、学習方法である。これらの違いは、3.1.2 で述べた通りである。

この方式では、単純 AE に基づく異常予兆検知と同様に、互いに関係するプロセス値間の相関関係の崩れが異常予兆として検知される。単純 AE とは異なり、複数時刻のプロセス値を扱うため、時間方向の相関関係が訓練データと異なる異常予兆も検知できる。

4. 評価結果

2段階 AE に基づく異常予兆検知方式の有効性を確認した評価の結果を示す。比較手法は、2.2 に示した単純 AE に基づく方式とした。

データとしては、稼働している発電所のボイラ系統のプロセス値を利用した。サンプリング間隔は1分とした。各 AE のモデルを学習するための訓練データ向けに、約 300,000 分間のデータを準備した。各方式を適用するテストデータとしては、訓練データとは期間が重ならず、時刻が後の約 10,000 分間のデータを準備した。

テストデータの期間には、ボイラ燃焼室のある位置に設置された温度計の測定値が実際よりも低くなる劣化が含まれていた。この劣化を各 AE で検知できるか否かを評価した。また、他のプロセス値の正常な挙動を誤って異常予兆として検知しないかどうかを評価した。

訓練データで学習した両 AE をテストデータに適用した結果、当該温度計について、図 1 0 に示す再構成値の相対値を得た。この図に、当該温度計の測定値も示す。ここで、横軸は、テストデータの最初の時刻を 0 とした相対時刻を表す。縦軸は、当該温度計の最初の時刻における測定値を 0 とした各値の相対値を表す。黒と薄い灰色がそれぞれ 2 段階 AE と単純 AE による再構成値の相対値を表し、濃い灰色が測定値の相対値を表す。

図 1 1 に、当該測定値の各 AE の再構成値からの偏差を示す。ここで、黒と薄い灰色が測定値の 2 段階 AE と単純 AE による再構成値からの偏差をそれぞれ表す。

図 1 0、1 1 から、測定値が各 AE の再構成値から乖離した様子が見てとれる。偏差の絶対値が温度計の測定誤差とはみさせない程度に大きかったことから、閾値を過大に設定しなければ、いずれの AE でも劣化を検知できることがわかる。

各 AE で閾値を変化させて、相対検知タイミングと相対誤検知率の関係をプロットしたグラフを図 1 2 に示す。ここで、横軸が相対検知タイミングを表し、縦軸が相対誤検知率を表す。相対検知タイミングは早いほど、相対誤検知率は低いほど好ましいことから、グラフにおける位置がより左下であるほど優れていることを表す。各方式の配色は、図 1 1 と同じである。

図 1 2 から、2 段階 AE の方が単純 AE よりも全体的に誤検知率が低かったことを確認できる。相対検知タイミングが 0 の場合では、単純 AE の相対誤検知率 100% に対し、2 段階 AE の相対誤検知率は 70.3% であった。

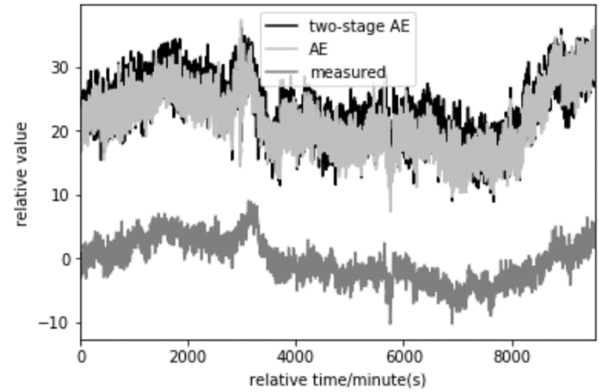


図 1 0 測定値および単純 AE と 2 段階 AE の再構成値の相対値

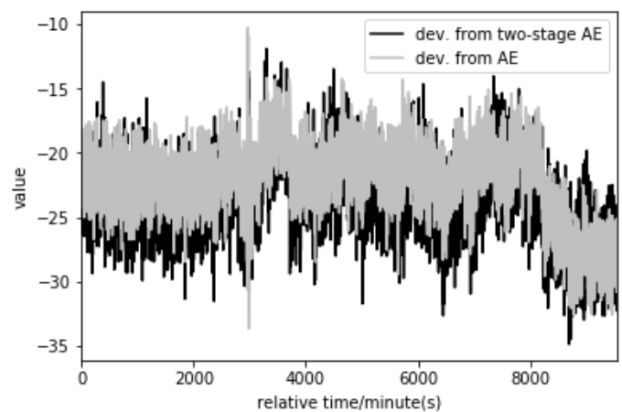


図 1 1 測定値の単純 AE と 2 段階 AE の再構成値からの偏差

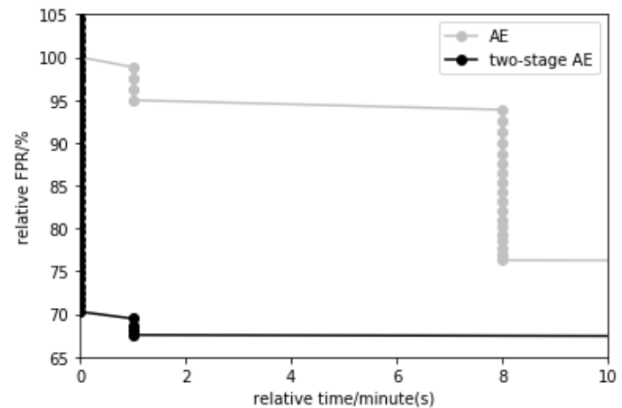


図 1 2 相対誤検知率

5. おわりに

稼働中発電所のボイラ系統のプロセスデータに 2 段階 AE を適用し、ボイラ燃焼室に設置された温度計の劣化の検知に成功した。この劣化は、関係するプロセス値との相関関係を崩すため、単純な AE でも検知されたものの、誤検知率が 2 段階 AE の方が低かった。これにより、2 段階 AE の実用性を示した。

参考文献

- [1] Guansong Pang, Chunhua Shen, Longbing Cao, and Anton Van Den Hengel, “Deep Learning for Anomaly Detection: A Review”, *Journal of ACM Computing Surveys*, Vol. 54, No. 2, 2021, Article No. 38.
- [2] G. E. Hinton and R. R. Salakhutdinov, “Reducing the Dimensionality of Data with Neural Networks,” *Science*, Vol. 313, No. 5786, 2006, pp.504-507.
- [3] Susumu Naito, Yasunori Taguchi, Yuichi Kato, Kouta Nakata, Ryota Miyake, Isaku Nagura, Shinya Tominaga, and Toshio Aoki, “A new data driven method for monitoring a large number of process values and detecting anomaly signs with a two-stage model composed of a time window autoencoder and a deviation autoencoder”, *Proceedings of the 2020 International Conference on Nuclear Engineering (ICONE2020)*, 2020, Paper No. ICONE2020-16150.
- [4] 名倉伊作、富永真哉、三宅亮太、青木俊夫、内藤晋、田口安則、加藤佑一、中田康太、 “A Iを用いた原子力発電所における異常予兆検知システムの開発 (1)オートエンコーダを適用した検知アルゴリズムの検討”、日本原子力学会 2020 年春の年会予稿集、2020、2J01.
- [5] 富永真哉、名倉伊作、三宅亮太、青木俊夫、内藤晋、田口安則、加藤佑一、中田康太、 “A Iを用いた原子力発電所における異常予兆検知システムの開発 (2)原子力発電所への導入に向けた構成検討”、日本原子力学会 2020 年春の年会予稿集、2020、2J02.
- [6] 名倉伊作、富永真哉、三宅亮太、青木俊夫、内藤晋、田口安則、加藤佑一、中田康太、 “A Iを用いた原子力発電所における異常予兆検知システムの開発 (3)異常予兆検知システムの全体概要”、日本原子力学会 2020 年秋の大会予稿集、2020、2K01.
- [7] 内藤晋、田口安則、加藤佑一、中田康太、名倉伊作、富永真哉、三宅亮太、青木俊夫、 “A Iを用いた原子力発電所における異常予兆検知システムの開発 (4) 2段階オートエンコーダによる検知アルゴリズムの改良”、日本原子力学会 2020 年秋の大会予稿集、2020、2K02.
- [8] 富永真哉、名倉伊作、三宅亮太、青木俊夫、内藤晋、田口安則、加藤佑一、中田康太、 “A Iを用いた原子力発電所における異常予兆検知システムの

開発 (5)原子力発電所への導入へ向けたユーザーインターフェースの検討”、日本原子力学会 2020 年秋の大会予稿集、2020、2K03.