

# LSTMに基づくノイズ除去オートエンコーダによる時系列予測

安藤大貴<sup>†</sup> 飯間等<sup>‡</sup>

京都工芸繊維大学 情報工学専攻<sup>†</sup> 同 情報工学・人間科学系<sup>‡</sup>

## 1. はじめに

オートエンコーダ(AE)は事前学習で入力の特徴を抽出するために用いられ、時系列の予測に対しては、Long short-term memory(LSTM)を組み合わせたLSTM-AEが提案されている[1]。しかし一般にAEには、入力をコピーして出力するように学習して、入力の特徴を適切に抽出できない場合がある。この問題を解決するために、ノイズ除去AE(Denoising AE; DAE)による学習を複数回行う積層型DAE(Stacked DAE; SDAE)が提案されている[2]。本稿では、時系列の予測性能をより高めるために、DAEをLSTMと組み合わせた事前学習法を提案する。

## 2. 積層型ノイズ除去オートエンコーダ

SDAEではDAEによる学習を複数回実行する。毎回の学習では、入力に小さなノイズを加えたものをDAEに入力し、ノイズを除去した入力を再構成する。各学習後にエンコーダを保存し、次の学習では新しいエンコーダとデコーダを連結して学習する。

## 3. 時系列予測の事前学習を行う提案手法

提案手法はSDAEとLSTMを組み合わせたものとする。すなわち、エンコーダとデコーダをLSTMとしたDAE(LSTM-DAEと呼ぶ)をモデルとし、このLSTM-DAEを用いた学習を複数回実行して事前学習を行

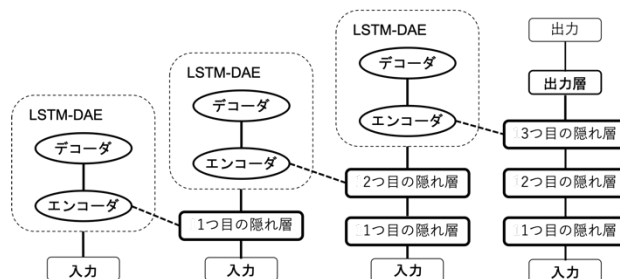


図1. 提案手法による事前学習の例

う。図1は、3回実行する場合の事前学習を示している。

事前学習の手順は、以下の3ステップからなる。

1. 最初のLSTM-DAEブロックを用いて学習し、そのエンコーダを1つ目の隠れ層として保存する。
2. 保存した隠れ層に新しいLSTM-DAEブロックを連結して学習させる。ここでは新たに連結したLSTM-DAEのパラメータのみを学習し、保存した隠れ層のパラメータの値は変更しない。また、元の入力の特徴を学習させるために、隠れ層から出力されるものを再構成するのではなく、元の入力を再構成するように学習させる。そして、ノイズは隠れ層から出力されるものに加える。学習後のLSTM-DAEのエンコーダを保存する。
3. 2を必要回数繰り返す。

ノイズには、通常のDAEで用いられるガウシアンノイズ(gaussian)、0に変更するマスキングノイズ(masking0)に加えて、以下のノイズも用いる。

- 同じ種類の時系列データの中で、時間の異なるデータを入れ替えるスワップ

Time Series Prediction by an LSTM-based Denoising Autoencoder  
 Daiki Ando, Information Science, Kyoto Institute of Technology  
 Hitoshi Iima, Information and Human Sciences, Kyoto Institute of Technology

ノイズ(swap)

- スワップノイズの入れ替える範囲をある時間以内に限定したノイズ(swap24)
- 連続する 2 つの時刻のデータを両方とも 0 に変更するマスキングベルトノイズ(masking0belt2)
- 平均値に変更するマスキングノイズ(maskingavg)
- 最小値に変更するマスキングノイズ(maskingmin)

以上に提案した事前学習を行った後は、時系列を予測するための学習を行う。

#### 4. 実験

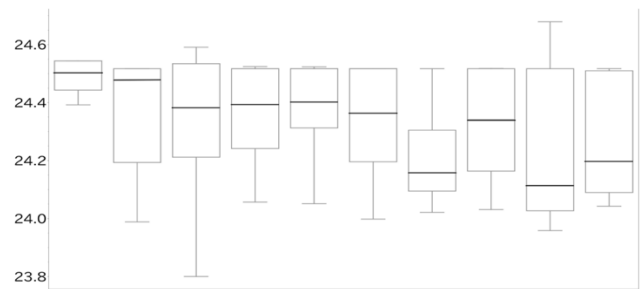
中国の北京市における毎時間の PM2.5 濃度を予測する[3]。入力時系列は PM2.5 濃度、露点、気温、気圧、風向、風速、累計積雪時間、累計降雨時間の 8 種類である。

ノイズに関して、gaussian は全てのデータに加え、それ以外のノイズはデータ毎に確率0.1で加える。swap24 の時間範囲は 24 時間以内とする。さらに、masking0 と masking0belt2 に対しては、データの正規化の前(before)と後(after)でノイズを加えるそれぞれ 2 種類の実験を行う。

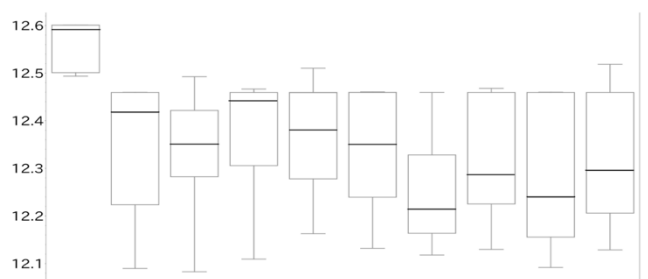
各手法による学習を各 20 回行ったときの 3 種類の誤差を図 2 に示す。この図において、左端は既存手法 LSTM-AE の誤差、その右以降は加えたノイズごとの提案手法 LSTM-SDAE の誤差を表す。図 2 より、各誤差の平均値については、LSTM-SDAE が LSTM-AE より小さくなっていることがわかる。特に、masking0\_before では、ばらつきを考慮しても、3 つ全ての誤差が低いので優れた性能を有している。以上より、LSTM-SDAE は有用であると言える。

#### 参考文献

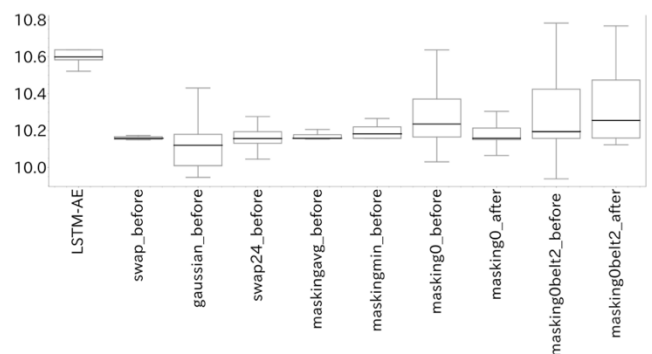
[1] A. Sagheer and M. Kotb, "Unsupervised pre-training of a deep LSTM-based stacked autoencoder for multivariate time series forecasting problems," Scientific reports, vol.9, 19038, 2019.



(a)二乗平均平方根誤差



(b)平方絶対誤差



(c)対称平均絶対誤差

図 2. PM2.5 予測誤差

[2] P. Vincent, H. Larochelle, et al., "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," Journal of Machine Learning Research, vol.11, no.110, pp.3371-3408, 2010.

[3] X. Liang, T. Zou, et al., "Assessing Beijing's PM2.5 pollution: severity, weather impact, APEC and winter heating," Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol.471, no.2182, 2015.