

## 線形回帰モデルを活用した深層時系列予測モデルの比較検討

## Comparison of Deep Time Series Prediction Models Utilizing Linear Regression Model

保住 純<sup>\*1</sup> 松尾 豊<sup>\*1</sup>

Jun Hozumi Yutaka Matsuo

<sup>\*1</sup>東京大学

The University of Tokyo

Time series prediction is important for the industry and various prediction models have been proposed. Recently some reports showed that deep learning models have higher accuracy than traditional models. However, it is known that these models take longer time to learn and are difficult to maintain long-term periodicity. To deal with these problems, some researches on deep models with the concepts of autoregressive models have been proposed. However, since these models have not been compared with the same data and same settings, it is unclear which models are effective for different tasks and different data. In this research, we apply these models to different types of time series data under different settings and examine the learning processes and results to capture the features of each model. The results confirmed the claimed merit of the model and suggested that simpler deep models are effective under complicated problems.

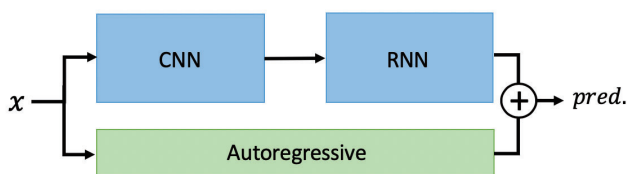


図 1: LSTNet の概念図

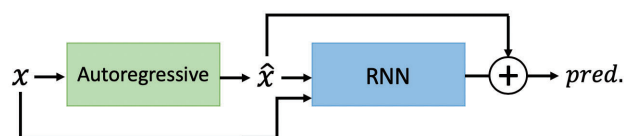


図 2: R2N2 の概念図

## 1. はじめに

時系列データの将来予測は産業上非常に重要であるため、これまでに様々な予測モデルが提案されてきた。また、機械学習分野における予測モデルの研究では、深層学習によるモデルが従来のモデルに比べて高い精度が出る事例も報告されている。しかし、深層学習を用いた予測モデルは、データ量が多くなると学習に要する時間が増大することや、長期にわたる周期性を保持し続けることが難しいことから、そのような性質を持つ時系列の予測には不向きであることも知られている [Goel 17]。そして、株価や商品売上など人間活動の記録によって発生する時系列データは、一週間や一年間といった複数種類の周期による影響を受けることが考えられる。

以上の問題を踏まえ、深層学習を用いたモデルの研究では、これまでに自己回帰モデル (Autoregressive Model; AR) の発想を取り入れることで周期性の保持や計算量の削減を実現した複数種類の手法 [Bińkowski 17, Goel 17, Lai 18] が提案されてきた。しかし、これらのモデルは同じデータや設定での比較した研究はなされていないため、データごとの各タスクに対しどのモデルが有効であるかは不明な点も残されている。

そこで、本研究ではこれらの手法を複数種の時系列データに対し異なる設定下で適用し、結果を比較検討することで各モデルの特徴を捉えることを目的とする。この成果は実業務における時系列予測において有用な示唆を示すとともに、今後の時系列予測手法の研究にも活かせるものであると考えられる。

連絡先: 保住 純, 東京大学, hozumi@weblab.t.u-tokyo.ac.jp

## 2. 前提知識

本研究における予測の対象は一定の時間間隔でサンプリングされた時系列を想定し、以降時刻  $t$  において観測された変数  $x$  を  $x(t)$  と記す。

## 2.1 線形自己回帰モデル

オーダーが  $N$  である線形自己回帰モデル  $AR(N)$  は以下の式で示される。

$$x(t) = \sum_{k=t-(N-1)}^{t-1} a_k x_k + c + \epsilon(t) \quad (1)$$

式 1 中の  $c$  は定数項,  $\epsilon$  はノイズを表す。

また、深層ネットワークにおいて AR モデルはベクトル  $\mathbf{x}_{[t-k:t-1]}$  の単純な全結合層によって、ノイズ部分以外を計算することができる。

## 2.2 線形回帰モデルを活用した深層モデル

本節では前述した深層モデルの構造の概要および線形回帰モデルとの関連を紹介する。詳細については各論文を参照されたい。

## 2.2.1 LSTNet

LSTNet(Long- and Short- term Time-series Network)[Lai 18] は、図 1 で表されるように、深層ネットワークによる予測と線形自己回帰モデルによる予測を並行して実行し、最後に両方の出力の和をとって最終出力とするモデルである。深層ネットワーク側は CNN 層, RNN

層, 全結合層の順で構成される. [Lai 18] ではタスクに応じて Skip-connection や Attention Layer を導入することで精度を高める試みもなされている.

### 2.2.2 R2N2

R2N2(Residual RNN)[Goel 17] は, 図 2.2 で表されるように,  $x(t-1)$  までを入力として  $x(t)$  を線形自己回帰モデルで予測を行い, その後, 真の  $x(t)$  の値との差分を RNN によって予測し, 最後にそれらの和を予測値とするモデルである. 言い換えるなら, R2N2 は線形自己回帰モデルを用いて大部分の予測を実行し, その結果生じた残差  $\epsilon(t)$  に相当する部分のみを RNN で学習させることで精度を高めるモデルであると言える.

### 2.2.3 SOCNN

SOCNN(Significance-Offset Convolutional Neural Network)[Bińkowski 17] は, 時間方向の畳込み (Significant-Network) によって式 1 における  $a$  の学習, データ間の畳込み (Offset-Network) によって  $x$  の補正を行い, 最終的にそれらを式 1 に当てはめて出力を算出するモデルである. Offset-Network は入力が多変量である時系列に対して同時刻  $t$  における特徴量の畳み込みを行うため, より強い補正効果が期待される.

## 3. 実験

### 3.1 設定

本研究では, 前章で紹介した LSTNet, R2N2, SOCNN の他に, 比較対象として線形回帰 (Linear), 3 層の多層パーセプトロン (MLP), 2 層 CNN, 2 層 LSTM による時系列予測実験を行った. これらのモデルは容量に大きな差が出ないよう層の数やノードの数を調整している. 活性化関数は LeakyRelu, 最適化関数は Adam, 損失関数及び評価指標は平均二乗誤差 (MSE) として, それぞれ 5000epoch 学習させた. また, いずれのモデルも勾配が最大 1.0 になるよう Gradient Clipping がなされている.

また, 各予測モデルの特長を捉えるために, 2 種類の時系列データで実験を実施した. 1 つ目は人工データとして, 以下の式 2 で表される周期性を持った時系列データを  $t = 0 \sim 1825$  の区間で作成した.

$$x(t) = \sin\left(\frac{\pi}{7}t\right) + 2\sin\left(\frac{\pi}{30}t\right) + 4\sin\left(\frac{\pi}{365}t\right) + 0.005t + \epsilon(t) \quad (2)$$

この時系列は複数の周期と振幅を持つ波形と, 時刻とともに上昇するトレンドを足し合わせた波形となっており,  $\epsilon(t)$  は平均 0, 分散 1 の正規分布からなるノイズ項である. 2 つ目は実世界のデータとして, 日毎のドル円為替データを 5 年分<sup>\*1</sup>を使用した.

どの実験においても入力データ長は 90 とし, 訓練データとテストデータはレコード数が 4:1 になるよう時間方向で分割し, 双方に共通の変数が含まれないようにしてある. また, 予測対象を入力データの直後 (0 日後) および 60 日のスキップ期間後という, 複数の設定下で実験を行った.

### 3.2 結果

実験結果を表 1 に記す. 括弧内の数字は入力データから予測対象までの日数を示す.

表 1: 各モデルの予測最高スコア (MSE)

モデル	人工 (0)	人工 (60)	為替 (0)	為替 (60)
Linear	10.69	10.62	8.66	10.64
MLP	11.03	10.83	8.24	17.39
LSTM	10.35	11.21	121.32	126.12
CNN	11.69	10.89	12.75	<b>6.21</b>
LSTNet	10.57	10.44	8.43	10.59
R2N2	10.62	<b>10.39</b>	8.47	10.58
SOCNN	<b>10.00</b>	10.52	<b>4.54</b>	6.44

まず, 周期性を持たせた人工データの予測では, 線形回帰モデルを併用したモデルはより高い精度を示しているものの, 他の単純なモデルにおいてもそれに近い誤差を記録している. また, 人工データの直後予測実験における, Epoch 数ごとのテスト誤差の推移を図 3 に示す. Epoch ごとの誤差の変化に注目すると, LSTNet, R2N2, SOCNN といった線形回帰を活用したモデルが DNN や LSTM など他の深層モデルに比べて早期に収束していることが確認される. このことから, 時系列予測における深層モデルは, これらの論文で主張されている通り, 線形モデルの併用が学習に要する時間を減らす効果があることが示唆される. また, この傾向は後述するドル円為替の時系列予測においても確認されている.

次に実データとしてのドル円為替の予測の結果を見ると, 線形モデルを併用した深層モデルが全体的に良い精度を示している. また, 直後については SOCNN, 60 日後においては CNN が最も良い精度を示している. CNN を活用したモデルが高い精度を示している理由としては, 畳込みによって時系列に含まれていたノイズの影響が軽減されていることにあると考えることができる. また, 時間間隔を開けた場合には単純な CNN が精度を上回る結果となっていることや, LSTM ではどちらの実験においても良い精度が出ずに人工データの直後予測以外では他のモデルに比べて精度が悪くなっていることを踏まえると, 予測対象までのスキップ期間が空いてしまう時系列予測タスクにおいては, 線形モデルの併用や RNN の採用が予測に悪影響を及ぼしている可能性が示唆される.

以上の考察を踏まえると, そもそも自己回帰モデルが有効であると考えられる時系列予測タスクについては, 線形モデルを活用した深層モデルが高い精度を示すものの, 周期性が不明確であったり予測対象にスキップ期間が挟まるような複雑なタスクでは, 単純な深層モデルや, 線形モデルや RNN を採用せずに CNN を重視するモデルを用いるほうが, より高い性能を発揮されることが考えられる.

## 4. まとめ

複数のモデルで実験を行った結果, 線形自己回帰モデルが有効であると考えられる時系列予測タスクには, 線形回帰モデルを活用した深層モデルがより有効であると言えるが, そう言えない時系列予測タスクはより単純な深層モデルや CNN を活用したモデル有効であることが示唆された. この結果は株価や消費電力量といった異なる実データや, スキップ期間の幅を変更する等さらなる検証実験が必要であると考えられる.

また, 本研究では入力をスカラー列として実験したが, 今回参考としたモデルは入力がベクトル列である多変量時系列回帰問題を想定しているため, その設定下では結果が変わることも大いに想定される. その場合における検証も今後の課題とする.

\*1 <https://www.federalreserve.gov/releases/h10/Hist/>

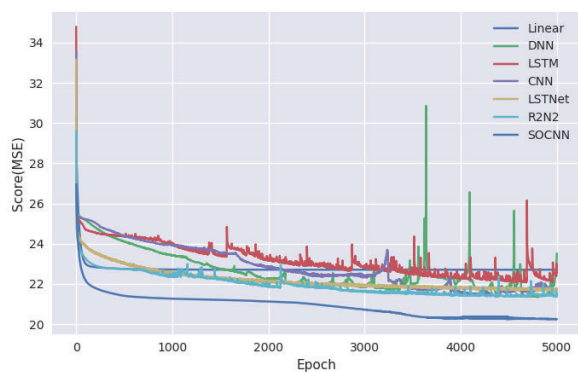


図 3: 人工 (0) における各 Epoch ごとのモデルのテスト誤差

## 参考文献

- [Bińkowski 17] Bińkowski, Mikołaj, Gautier Marti, and Philippe Donnat. "Autoregressive convolutional neural networks for asynchronous time series." arXiv preprint arXiv:1703.04122 (2017).
- [Goel 17] Goel, Hardik, Igor Melnyk, and Arindam Banerjee. "R2N2: Residual recurrent neural networks for multivariate time series forecasting." arXiv preprint arXiv:1709.03159 (2017).
- [Lai 18] Lai, Guokun, et al. "Modeling long-and short-term temporal patterns with deep neural networks." The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. ACM, 2018.