6U-01

Transformer を用いた COVID-19 の感染者数の推移予測におけるノイズ付加の効果分析

山下穂乃花†　　藤田悟‡

法政大学　情報科学部†

1. まえがき

COVID-19 は、現在も感染拡大が止まらず、多数の感染者数が報道されている。本研究では、COVID-19 の医療対策などを円滑に行うために役立つよう、感染者数予測を精緻化する。具体的には、感染者数の推移を、Transformer で学習し、未来の感染者数を予測することを目的とする。課題として、COVID-19 は約 3 年前から感染が始まり、ニューラルネットワークの学習を行うには、データ数が限られている。そこで、微少ノイズを付加してデータ増強を行い、予測精度の向上を目指す。本稿では、ノイズの付加方法の提案と、ノイズ付加をするときとしないときの感染者数の予測性能への影響を分析する。

2. Transformer

Transformer とは、Google が開発した機械翻訳のための深層学習モデルである。同じ時系列予測モデルである LSTM とは違って再帰性を排除し、入力と出力の依存関係を類似度行列で描く Attention メカニズムによって学習できる[1]。

Transformer は、機械翻訳を目的としているため、原語を入力する encoder と対象言語を出力する decoder から構成されている。Transformer の decoder は、対象言語の時系列入力の先読みが出来ないように masking 機能が実装されている。本稿では、1 次元の入力データを学習するため、学習モデルとして Transformer の decoder を利用して、未来予測モデルを構築した。

3. COVID-19 の感染者数の予測方法

3.1. COVID-19 感染者数のデータセット

本稿で扱うデータは、厚労省のホームページに掲載されている「新規陽性者数の推移」というオープンデータである[2]。このデータセットは、都道府県ごとに、2020 年 1 月 16 日から現在まで 1 日ごとの新規陽性者数が記載されている。本稿では、予測を正確に行うため、最初の感染が確認された 2020 年 1 月 16 日から感染者数が少ない期間である 200 日を除き、2020 年 8 月 3 日からの東京都の新規陽性者数のデータのみ使用する。

3.2. 予測方法

本稿では、感染者数のダイナミックレンジを縮め、予測しやすくするために、感染者数の対数を取り、0 から 1 の範囲に正規化したデータを扱う。本稿ではt日目から$t+i$日目までの入力$X_{t,i}$を以下のような形で扱う。ここで、x_{t+i}は、$t+i$日後の感染者数の対数とする。

$$X_{t,i} = (x_t, x_{t+1}, x_{t+2}, \ldots, x_{t+i}\})$$

セット数Tの入力$(X_{1,i}, X_{2,i}, \ldots, X_{T,i})$に正規分布$N(0, 0.001)$に従ったノイズを追加した入力を、作成した学習モデルに入力し、学習させる。　a日先までの予測をするとしたとき、入力$(X_{1,i}, X_{2,i}, \ldots, X_{T_1,i}\})$を学習した後の出力は、$(X_{1+a,i}, X_{2+a,i}, \ldots, X_{T_2+a,i})$となる。

$(X_{T+1,i}, X_{T+2}, \ldots, X_{T_2+a,i})$を予測結果として用いる。

3.3. 学習モデル

本稿では Transformer の decoder を用いて学習モデルを作成した。作成した学習モデルを図 1、図 2 に示す。

図 1 に示すように、入力$(X_{1,i}, X_{2,i}, \ldots, X_{T,i})$を受け取った後に Dense 層を置いて次元を拡張した。これは、感染者数の各時刻のデータは 1 個の実数値であるため、ニューラルネットワークの学習に適さないからである。データの次元拡張により、例えば 128 ノードに展開でき、以降のニューラルネットワークの学習が可能になる。Dense 層を 2 層置いた後、PositionalEncoding 層で位置情報を付加し、TransformerDecoder で学習を行う。後段として、入力段の逆変換を行うために Dense 層を 2 層置き、拡張した次元をもとに戻す。

図 2 は、Transformer の前後を skip connection で結ぶことで、前日との感染者数との差分を中心に Transformer で学習するモデルに改良したものである。前日との感染者数との差分を学習する以外の部分は、図 1 と同じように Dense 層を配置した。

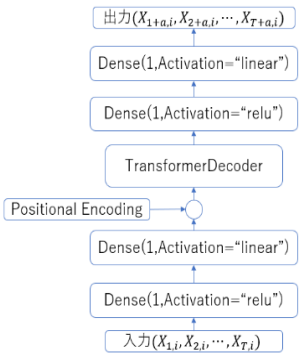

図 1. 作成した学習モデル

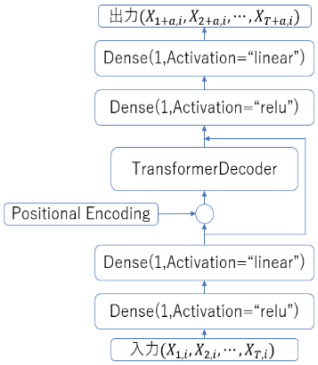

図 2. skip connection を利用した学習モデル

Analysis of the Effect of Noise Addition on the Prediction of the Number of COVID-19 Infections Using Transformer

† Honoka Yamashita, Satoru Fujita, Computer and information Science, Hosei University

Copyright ©2023 Information Processing Society of Japan.
All Rights Reserved.

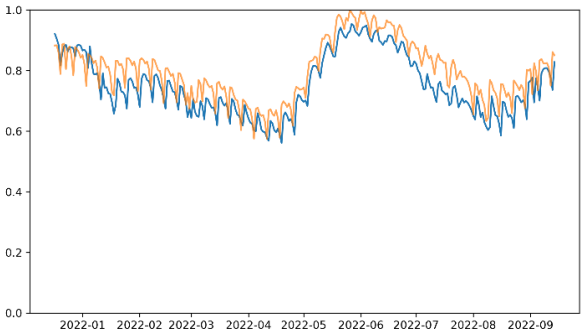

図 3

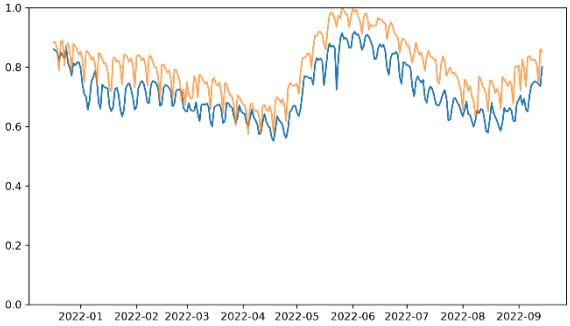

図 4

図 5

4. 実験

4.1. 実験方法

実験では、学習対象のデータ長を 100 日分とし、それを 1 日ずつずらして作成した 500 セットの訓練データを用意し、さらに、130 セットのテストデータを用意した。入力データの次元拡張は 256 ノードである。予測するデータは翌日の感染者数とした。

Skip connection を利用しない図 1 の学習モデルを実験 1 とし、skip connection を利用する図 2 の学習モデルを実験 2 とした。この両者を比較することで、skip connection の有効性を評価する。

さらに、3.2 節に述べたデータにノイズを加えた効果を比較するために、ノイズを加えない学習を行ったモデルを実験 3 とする。学習モデルは、図 1 の skip connection のないモデルを用いた。

4.2. 実験結果

実験 1, 実験 2 の学習後の予測結果を図 3, 図 4 に示す。予測値を青線、正解値をオレンジ線で示している。実験結果のグラフは、130 セットのテストデータの予測結果を結んだものであり、130 日分を一度に予測したものではないことに注意する。すなわち、常に、1 日前までの感染者数を知った上で、翌日の予測を行ったものである。

図 3 より、感染者数が増加、減少傾向にある予測のときに、予測が多きくはずれていることがわかる。一方で、1 週間の曜日による感染者数の変化の様子は予測できていることが読み取れる。図 4 より、常に正解値より低い数値を予測してしまっていることがわかる。

次に、図 3 と図 4 のときの実験結果の精度比較をした。その結果を表 1 に示す。 表 1 より、skip connection を用いたときのモデルでの学習の方が、全体的に予測精度が低いことがわかる。

図 1 の学習モデルを使用してノイズを付加しないときの翌日の予測を行った。結果を図 5 に示す。図 5 より、予測が大きく外れていて 1 週間の中での変化の予測にも対応できていないことがわかる。ノイズを付加したときと付加しないときの精度の比較を表 2 に示す。表 2 より、ノイズを付加した方が、精度が少し良くなっている。

5. 考察

翌日の予測について、1 週間の中での感染者数の増減は予測できていたが、全体的には正解値より低い数値が予測されていた。理由としては、学習モデルに改良の余地がある、2022 年以前で高い感染者数の数値が継続したことがあまりなかったことなどが考えられる。また、skip connection での予測精度が低くなった理由としては、感染者数を 0 から 1 の範囲に正規化したため、差分の値が小さく、認識が難しかったことなどが考えられる。

6. むすび

本稿では Transformer を用いて新規コロナ感染者数の予測を行った。1 週間の中での感染者数の変化は予測できていたが、1 日ごとの細かい予測の精度をあげることができなかった。学習モデルの構築方法を見直す、Transformer の decoder のパラメータ等を細かく調節することにより、1 日後の予測精度を上げる必要がある。

参考文献

[1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin, "Attention is all you need", 2022/11/01

[2] オープンデータ | 厚生労働省, https://www.mhlw.go.jp/stf/covid-19/open-data.html, 2022/11/01

表 1. 図 3 と図 4 の実験結果の比較

	図1のモデルを使用したとき	図2のモデルを使用したとき
MAE	0.0424	0.0521
RMSE	0.0529	0.0626

表 2. ノイズによる効果比較

	ノイズあり	ノイズなし
MAE	0.0424	0.0449
RMSE	0.0529	0.0553

Copyright ©2023 Information Processing Society of Japan.
All Rights Reserved.