

## 深層学習による翌日電力需要予測手法の検討 Study of Daily Power Consumption Prediction Using Neural Network

近藤 海夏斗<sup>†</sup> 内保 光太郎<sup>‡</sup> 西山 英治<sup>†</sup> 新谷 洋人<sup>†</sup>  
Minato Kondo Kotaro Uchiho Eiji Nishiyama Hirohito Shintani

### 1. はじめに

需要予測とは、商品やサービスに関する将来の需要量を予測することである。電力をはじめとするインフラ分野において、需要予測は運用計画の起点となる。安定した電力供給を実現するためには、需給の一致が必要不可欠であり、一般送配電事業者が日々調整をおこなっている[1]。電力需要は、主に再生可能エネルギー発電の発電出力量や気象状況、そして曜日や時間帯といった項目を主要因として変動することが知られている[2, 3]。そのため、外乱要素が多く、電力需要に対するノウハウがなければ、予測精度を高い水準にすることは困難である。

また、現在発電事業者及び小売電気事業者は、毎日翌日の発電・需要の計画を電力広域的運用推進機関(OCCTO)へ提出している。これらの計画と実績の差をインバランスといい、インバランスに比例した金額を事後に精算している[4]。これらのことから、需要予測の精度は、電力システムの安定性に繋がるほか、収益に影響を与える要因となり、その重要性を増している。

本研究では、予測手法として LSTM(Long Short Term Memory)[5]や GRU(Gated Recurrent Unit)[6]、及び Transformer[7]をベースとした深層学習モデルを使用し、予測精度の向上を目指すことを目的とする。具体的には、直近のデータだけでなく、過去のデータから予測対象と似た時系列パターンを抽出する機構を導入する。そして、直近のデータと抽出したデータを連結し入力することで、予測精度の向上を図る。本稿では、国内の電力需要実績データ、気象データをデータセットとして、実際の運用を想定し、翌日 24 時間の電力需要予測を行い、各手法の比較・検討を行った。

### 2. 関連研究

電力需要をはじめとする時系列情報予測問題では、深層学習を使ったモデルが多く提案されている。モデルの多くは、時系列データを再帰的に計算し、時間依存性を記憶する RNN(Recurrent Neural Network)が用いられてきた。RNN ベースのモデルは時間依存性を捉える一方で、複数の変数間に生じる局所的な依存関係や、時系列の長期依存性を考慮していない。Lai 氏らはこれらの問題に対し、CNN(Convolutional Neural Network)、GRU、そして線形自己回帰を組み合わせた LSTNet(Long and Short term Time series Network)を提案した。LSTNet は、長期及び短期の時系列パターンが混在しているデータに対し優れた予測精度を示した[8]。

一方、自然言語処理では Transformer が提案され、機械翻訳をはじめとするタスクで優れた成果を挙げている[6]。Transformer は、RNN のような再帰的な処理を行わず、

Attention 機構をベースとすることで、高速化と長期依存性の学習を実現した。その汎用性の良さから、Transformer は時系列情報予測に応用した事例が報告されている[9]。Zhou 氏らは、Transformer が長期的な系列の予測に不向きであるという欠点を指摘し、この課題に対して Informer を提案した[10]。電力消費データ、気象データ等を用いて検証を行った結果、Informer は多くのデータセットおよび予測先において既存手法を上回る精度を達成した。Informer は以下 3 つの特徴を持つ。

(1) Self-Attention は系列長  $L$  に対して  $O(L^2)$  の計算量がかかる。そこで、ProbSparse Self-Attention 手法により、入力する系列長を近似計算により削減することで、計算量を  $O(L \log L)$  に抑えた。

(2) Self-Attention 層を出る度に、Conv1d と MaxPooling1d を通すことで系列長を圧縮した。

(3) 生成的デコーダーにより時系列を一度に予測し、推論速度を向上させた。

また、梶屋氏らは、Transformer に入力する系列長を減らす試みとして、Transformer をベースとしたモデルに、抽出モジュールを追加したモデルを提案した[11]。抽出モジュールとは、1 日前や 1 週間前などの過去系列から、予測時点付近と似た時系列パターンを探索し、抽出する機構である。人流予測について検証を行った結果、従来手法を上回る予測精度を示した。

### 3. 翌日電力需要予測の問題設定

本研究では、実際の運用を想定し、1 週間の電力需要、及び気温データから、翌日 24 時間の需要予測を出力するよう学習を行った。図 1 は東京電力エリアの電力需要実績データをもとに、問題設定の様子をグラフにしたものである。横軸は日付、縦軸は電力需要実績値である。黒線で示す 2016 年 4 月 1 日 0 時から 4 月 7 日 23 時の 7 日がモデルへの入力となる。そして、赤線で示す 4 月 9 日の 24 時間が予測対象、すなわちモデルの出力となる。

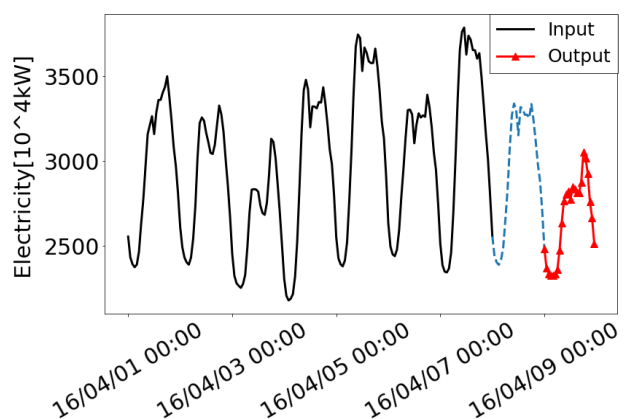


図 1 問題設定の様子

<sup>†</sup> 熊本高等専門学校 National Institute of Technology, Kumamoto College

<sup>‡</sup> 日本経済大学 Japan University of Economics

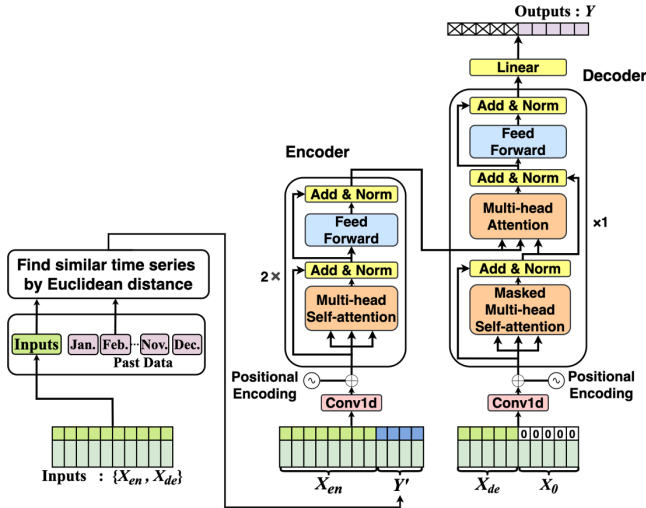


図 2 提案手法の概略図

#### 4. 提案手法

図 2 は、提案手法の概略図である。提案手法は、Transformer をベースとしている。Transformer の外部には、入力データから予測値に近い時系列パターンを検索する抽出モジュールが備わっている。

##### 4.1 抽出モジュール

抽出モジュールは、入力データによる予測対象の直近情報だけでなく、入力データと過去のデータから、予測対象に近い時系列パターンを探査することを目的としている。関連研究で述べた人流予測問題における抽出モジュール [11] は、検索対象を入力データから数日前までのデータを使用している。本研究では、電力需要が周期性に着目し、入力データから直近のデータに加え、1 年前のデータを使用する。図 3 は、東京電力の電力需要実績データにおける、四半期ごとの箱ひげ図である。横軸は四半期、縦軸は各四半期の電力需要実績値である。1 年単位で電力需要の分布に大きな差はないことが読み取れる。したがって、例えば 2 月の需要予測を行う際には、過去の 2 月のデータに、予測対象と似た時系列パターンが存在すると考えられる。以上のことから、本研究では、電力需要データに対し全期間での探索を行わず、

- ① 入力データの 2 週間前の区間
- ② 入力データの 1 年前から、前後 2 週間の区間の 2 区間で類似度を算出した。そして、入力データと最も類似していた部分の直後 48 時間分をエンコーダーの入力

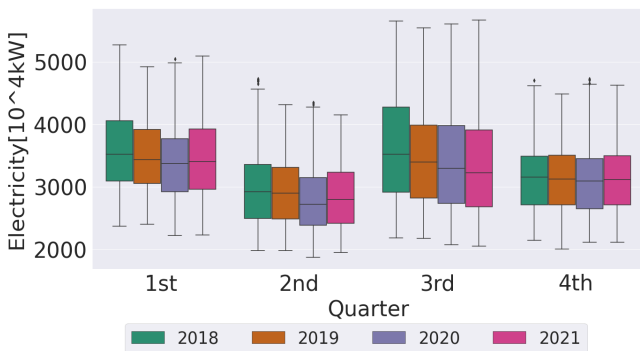


図 3 東京電力エリアの箱ひげ図

として使用した。例えば、2020 年 4 月 16 日を予測する際には、2020 年 4 月 8 日から 4 月 14 日入力データとなる。この時の検索区間は、

- ① 2020 年 3 月 25 日から 2020 年 4 月 7 日
  - ② 2019 年 3 月 25 日から 2019 年 4 月 28 日
- の 2 区間となる。なお、類似度は式(1)に示す、ユークリッド距離で算出した。

$$distance(x, t) = \sqrt{(x_1 - t_1)^2 + \dots + (x_n - t_n)^2} \quad (1)$$

#### 4.2 Transformer による時系列情報予測

Transformer は Attention 機構を用いた Encoder-Decoder モデルである。Encoder では、入力されるデータに Positional Encoding で位置情報を付与した後、Attention 機構により特徴量ベクトルを生成する。Decoder では、Encoder にて生成した特徴量ベクトルをもとに、予測値を生成する。本研究では、Encoder への入力を、入力データのうち最初から 6 日分と、抽出モジュールから出力された 2 日分のデータを連結したデータとした。

一方、Decoder は Informer で提案された生成的デコーダー [9] を用いた。本研究では、入力データの最後の 1 日と、予測対象の 24 時間分を 0 で masking したデータを連結し、入力する。出力では、0 で masking した箇所に、予測値が出力されるよう学習を行った。また推論時には予測値を一つずつ再帰的に出力せず、学習時と同様に masking した箇所の出力を予測値とする。また、活性化関数は ReLU、Encoder は 2 層、Decoder は 1 層、そして Multi-head Attention の head 数を 8 とした。

#### 5. 検証方法

##### 5.1 データセット

本研究では、表 1 に示す地域及び気温データを、それぞれ電力会社、気象庁のホームページから取得した。なお、気温データの地点は、経済産業省が公開している都道府県ごとの需要実績をもとに、各エリアで最も需要が高い都道府県の県庁所在地を選択した [12]。取得したデータは 1 時間ごとに抽出し、データ点数はそれぞれ 50424 点である。データの前処理として、式(2)を用いて最小値を 0、最大値を 1 となるよう正規化した。また、全データのうち、2016 年から 2019 年の 4 年分を学習データ、2020 年を検証データ、2021 年をテストデータとした。

$$Y = \frac{X - \min(X)}{\max(X) - \min(X)} \quad (2)$$

表 1 使用データ一覧

地域	気温データ対象都市	期間
東北電力	仙台市 (宮城県)	2016/4/1 0:00
東京電力	東京 (東京都)	
関西電力	大阪市 (大阪府)	~
中部電力	名古屋市 (愛知県)	2021/12/31 23:00
四国電力	松山市 (愛媛県)	
九州電力	福岡市 (福岡県)	

表 2 検証結果

モデル	評価指標	東北電力	東京電力	関西電力	中部電力	四国電力	九州電力
LSTM	MAE[万kW]	64.869	252.971	90.049	107.209	24.018	59.086
	MAPE[%]	6.707	7.343	5.280	7.107	7.635	5.996
GRU	MAE[万kW]	49.067	185.802	84.791	70.970	17.939	53.833
	MAPE[%]	4.930	5.504	4.999	4.590	5.582	5.355
LSTNet	MAE[万kW]	<b>46.230</b>	186.066	<b>79.214</b>	<b>69.210</b>	<b>15.959</b>	<b>50.087</b>
	MAPE[%]	<b>4.646</b>	5.504	<b>4.717</b>	<b>4.495</b>	<b>5.001</b>	<b>5.021</b>
Transformer	MAE[万kW]	51.650	196.916	83.285	79.473	17.352	56.065
	MAPE[%]	5.198	5.894	4.957	5.208	5.429	5.660
Informer	MAE[万kW]	79.799	190.464	98.802	111.614	22.355	64.354
	MAPE[%]	8.233	5.687	5.947	7.496	7.192	6.507
提案手法	MAE[万kW]	47.664	<b>178.524</b>	81.198	70.994	16.549	51.358
	MAPE[%]	4.830	<b>5.258</b>	4.801	4.601	5.207	5.105

## 5.2 比較モデル

提案手法に加え、GRU、LSTM、LSTNet、Transformer、そして Informer の 5 つを比較用として用いた。GRU、LSTM は、中間層に 2 層の GRU、もしくは LSTM、そして全結合層の 3 層で構成した。また、すべてのモデルにおいて、モデルの次元数を 128、256、512 の中から最も精度が優れていた値に設定し、10% の dropout を導入した。Transformer、及び Informer については、各種ハイパーパラメータを提案手法と同じ値とし、Informer の factor を 5 とした。

## 5.3 学習環境

学習を行うにあたって、最適化手法に Adam を使用し、パラメータの初期値は原論文で提案された数値に設定した[14]。損失関数は式(3)に示す、MSE(Mean Square Error)に設定した。検証データの MSE が収束するごとに、学習率を半減させて行き、再び収束した場合に学習を終了した。なお、テストデータでの推論時には、検証データで最も MSE が低くなった際のモデルの重みを使用した。さらに、学習の安定化のため、全モデルの学習時に、0.5 の勾配クリッピングを行った。実装には、深層学習ライブラリ Pytorch 1.1.0 を使用した。

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (3)$$

( $y$  : 実績値、 $\hat{y}$  : 予測値、 $N$  : データ数)

## 5.4 評価指標

予測精度の評価として、式(4)に示す MAE(Mean Absolute Error)、及び式(5)の MAPE(Mean Absolute Percentage Error)を用いた。なお、算出する際には、モデルの出力値を元のスケールに戻した。

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (4)$$

$$MAPE = \frac{100}{N} \sum_{i=1}^N |y_i - \hat{y}_i / y_i| \quad (5)$$

## 6. 検証結果

### 6.1 検証結果

表 2 に各モデル及びエリアごとの予測結果を示す。東京電力エリア以外において LSTNet が最も優れた精度を示した。LSTNet はニューラルネットワークによる予測値と線形自己回帰による予測値の和を最終出力としており、複数の学習器を用いるアンサンブル学習を取り入れたモデルとなっている。電力需要予測では、アンサンブル学習を取り入れることで高い予測精度を示す事例が報告されている[4, 15]。従って、LSTNet ではアンサンブル手法が予測精度の向上に寄与していると考えられる。

一方、提案手法は、すべてのエリアにおいて Transformer 及び Informer より優れた精度を示した。従って、電力需要予測においても過去の時系列を使用する手法は有効であることがわかった。

### 6.2 考察

図 4 はテストデータの 2021 年における、実績値とモデルの予測値をグラフにしたものである。(a)は 5 月 21 日の東京電力、(b)は 7 月 6 日の中部電力の予測値を示している。図 4 (a)より、提案手法が他のモデルに比べて、実績値と非常に近い値を出力していることが読み取れる。また、ユークリッド距離による検索結果も実績値とおおよそ近い値となっていることから、過去の類似している時系列パターンをもとに予測ができているといえる。

一方、図 4 (b)では、提案手法及び検索結果が実績値より大幅に低い値となっている。LSTNet 及び通常の Transformer では、提案手法より高い値となっているため、追加で入力した過去のデータの影響を受けて低い値となってしまったと考えられる。以上から、入力データに追加で入力した過去の実績値は、予測値に影響を及ぼしていることが確認できた。そして、予測値として出力すべき時系列パターンをより正しく抽出できれば、予測精度のさらなる向上が見込めることがわかった。

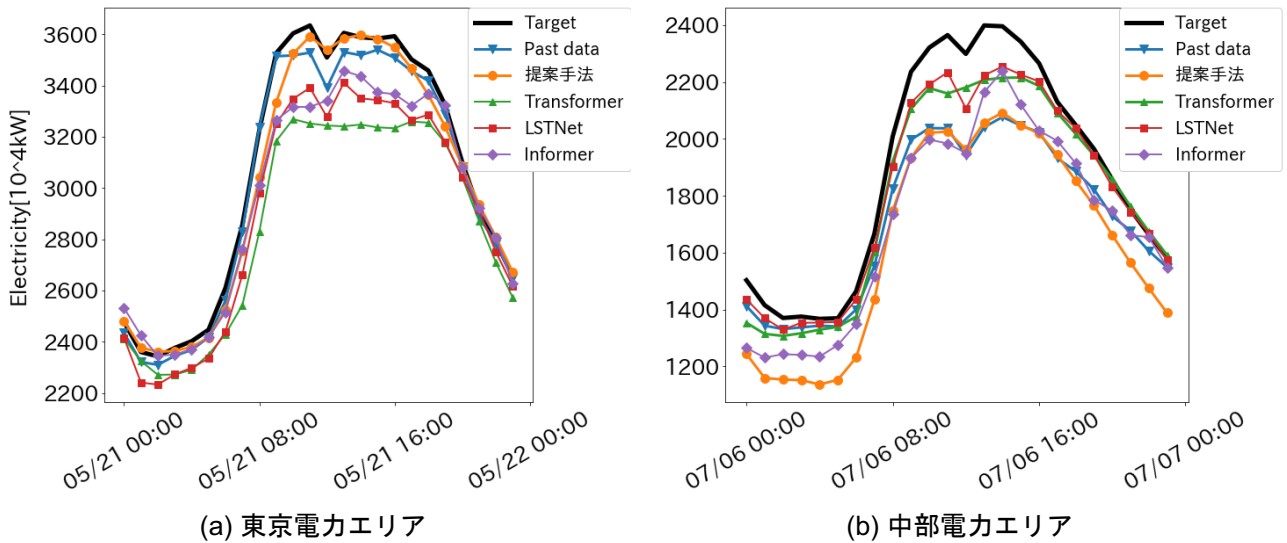


図 4 各モデルの予測結果

## 7. おわりに

本研究では、電力需要データがおおよそ 1 年の周期性があることに注目し、入力データの直近及び、1 年前のデータから類似した情報を抽出する機構を提案した。Transformer をベースとしたモデルに抽出モジュールを組み込み、国内の電力需要実績データ、気象データを用いて検証を行った結果、提案手法は通常の Transformer に比べ優れた精度を示した。また、RNN 及び Transformer をベースとしたモデルで検証を行った結果、東京電力を除くエリアで LSTNet が最も優れた精度を示した。このことから、予測精度の向上にアンサンブル学習は効果的であることがわかった。一方、予測結果の分析から、実際の値と抽出モジュールにより取得した過去のデータにずれがある箇所では、そのずれが原因で予測精度が悪化していることがわかった。今回は類似度指標として、電力需要データのみからユークリッド距離で算出しているため、気温や曜日などを考慮できていない。従って、複数の情報をもとに類似度を評価する手法の検討を行い、さらなる精度向上を目指す。

### 参考文献

- [1] 山口 頼央, “九州支部だより 九州のエネルギーと電気の「同時同量」について”, *電気設備学会誌*, Vol. 39, No. 10, pp. 661-662, (2019).
- [2] 竹田 恒, 岡本 和樹, 真保 崇, “電力需要予測コンテストオープンイノベーションへの取組み”, *人工知能*, Vol. 35, No. 3, pp. 360-365, (2020).
- [3] 猪股 海夏斗, 内保 光太郎, 西山 英治, 新谷 洋人, “気象データを用いた電力消費量予測精度の改善”, *第 20 回情報科学技術フォーラム*, (2021)
- [4] 進 博正, 志賀 慶明, 市川 量一, “気象予測データと機械学習を用いた高精度な電力需要予測手法”, *東芝レビュー*, Vol. 74, No. 5, pp. 22-25, (2019)
- [5] Hochreiter, S. and Schmidhuber, J., “Long short-term memory”, *Neural computation* 9(8): pp. 1735-1780, (1997).
- [6] Chung, J., Gulcehre, C., Cho, K., and Bengio, Y., “Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling”, *arXiv:1412.3555*, (2014).
- [7] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I., “Attention Is All You Need”, *arXiv:1706.03762*, (2017).
- [8] Lai, G., Chang, W.-C., Yang, Y., and Liu, H., “Modeling Long- and

Short-Term Temporal Patterns with Deep Neural Networks”, *arXiv:1703.07015*, (2018).

- [9] Li, S., Jin, X., Xuan, Y., Zhou, X., Chen, W., Wang, Y.-X., and Yan, X., “Enhancing the Locality and Breaking the Memory Bottleneck of Transformer on Time Series Forecasting”, *arXiv:1907.00235*, (2019)
- [10] Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., and Zhang, W., “Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting”, In *The Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Virtual Conference*, Vol. 35, pp. 11106-11115, AAAI Press, (2021).
- [11] 梶屋 友幹, 片山 晋, 浦野 健太, 青木 俊介, 米澤 拓郎, 河口 信夫, “Transformer モデルを用いた人流の時系列予測”, *マルチメディア、分散、協調とモバイル DICOMO2021 シンポジウム*, Vol. 2001, pp. 900-908, (2021)
- [12] Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”, *arXiv:1810.04805*, (2018)
- [13] 統計表一覧 | 資源エネルギー庁, [https://www.enecho.meti.go.jp/statistics/electric\\_power/ep002/result\\_s.html#headline2](https://www.enecho.meti.go.jp/statistics/electric_power/ep002/result_s.html#headline2), (参照日 2022 年 5 月 16 日)
- [14] Kingma, P. D., Ba, J., “Adam: A Method for Stochastic Optimization”, *arXiv:1412.6980v9*, (2017)
- [15] 柴田 智文, 大谷 公哉, 寄田 保夫, 川上 恭章, 永富 悠, 松尾 雄司, “人工ニューラルネットワークによる電力需要予測の影響要因評価 -学習条件による影響-”, *Journal of Japan Society of Energy and Resources*, Vol. 40, No. 5, pp. 144-153