

電力使用量予測のための深層学習手法における 最適なモデル選択に向けて

山名田 恭吾* 黒岩 丈介** 小高 知宏** 諏訪 いずみ*** 白井 治彦****

An Investigation of Optimal Model Selection in Deep Learning Methods for Electricity Usage Prediction

Kyogo YAMANADA*, Jousuke KUROIWA**, Tomohiro ODAKA**
Izumi SUWA*** and Haruhiko SHIRAI****

(Received February 1, 2023)

In this paper, we investigate deep learning methods to predict electricity consumption. We employ RNN (Recurrent Neural Network) and LSTM (Long Short-Term Memory), both of which are applicable in handling time-series data, to predict electricity usage. The target data are hourly, daily, and weekly electricity consumption data. The purpose of this paper is to investigate prediction accuracy for various structures of RNN and LSTM. The results show that the prediction accuracy surely depends on either RNN or LSTM and its structure for each of the hourly, daily, and weekly forecasts.

Key words :Electric Usage Prediction, Deep Learning, RNN, LSTM

1. はじめに

東日本大震災以降に、コストが低く、効率的に発電できる原子力発電が安全性に問題があると考えられている。原子力発電所の稼働の減少により、原子力発電よりもコストが高い火力発電を用いた発電が増えている。そのため、発電コストが増加し、電気料金の値上げが進んでおり、消費者の負担が増加している。近年、世界的に持続可能な開発目標 (Sustainable Development Goals) という活動が取り組まれている。2030年までに持続可能でよりよい世界を目指すとい

う目標を指しており、電力使用量の予測を行うことでコスト削減をはじめとするこれらのような取り組みに貢献ができるのではないかと考えた。

電力使用量の予測が可能になった場合には、長期的には季節ごとの電力使用量をもとに適切な時期の化石燃料の調達ができ、短期的には電力使用量が変動する時間帯に合わせて発電所の稼働状態を設定することができるようになる。これまでの電力使用量のデータから大まかな変動は予測が可能であるが、深層学習を用いて予測を行うことでより詳細に予測ができるようになり、適切な量の化石燃料の調達や、発電量の調節に役に立つ^[1]。適切な電力供給を行うことができれば、非効率な発電などを行う必要がなくなり、消費者の負担軽減にも繋がる^[2]。

人工知能に関する多くの分野で知的処理技術である深層学習が用いられている。深層学習は機械学習の一つの手法であり、ニューロンが多層に重ねられたニューラルネットワークを用いた分析を行うことである^[3]。深層学習は、表情認識や物体認識などの画像認識や、普段人間が用いる自然言語に対して解析する自然言語処理、音声で家電などを操作すること

*大学院工学研究科 知識社会基礎工学専攻

**Fundamental Engineering for Knowledge-Based Society, Graduate School of Engineering

***知能システム工学講座

****Department of Human and Artificial Intelligent Systems

*****仁愛女子短期大学 生活科学学科

*****Jin-ai Women's College

*****工学部 技術部

*****Technical Division

ができる音声認識などがある。深層学習は株価の予測や電力使用量の予測にも用いることができる。近年は深層学習が盛んに行われているが、課題もある。学習の時にデータ数が多いほど精度が高くなる傾向があるが、計算時間が長くなってしまったり、データが不足している場合には精度の向上が望めないことも挙げられる。また、ネットワークの構造によっても精度が異なってくることもあるため、複数のモデル構築を試したり、ハイパラメータを手動で変更し予測を行う必要があるなどの問題も存在する。人間の脳では、一つの情報からではなく、周りの状況などの他の情報からも情報処理が行われている。そのような処理の方法を模した手法にマルチモーダル学習がある。マルチモーダル学習は複数の情報を入力として学習を行う手法のことである。電力使用量が増える要因は気温や天気、降水量など様々考えられる。そこで、電力使用量だけではなく、他要因も組み合わせて入力として、学習することで予測精度が高くなることが考えられる。

本研究の目的は、電力使用量の時間変化量のみから電力使用量を予測する最適なモデル及びモデル構造を明らかにすることである。これにより、電力使用量が突然跳ね上がるような異常値の予測を実現したり、マルチモーダル入力による高精度な電力使用量予測を実現する研究へ繋げていくことを目的としている。

2. 時系列データ処理のための深層学習手法

2.1 ニューラルネットワーク

ニューラルネットワークは、人間の神経回路網をもとにして作られた数学モデルのことである。人間の脳は情報伝達や記憶の定着のために神経細胞の結びつきであるニューロンを用いており、脳神経系の強力な学習能力を画像認識などの分野へ応用する研究が近年盛んに取り組まれている。ニューラルネットワークは、入力層と中間層、出力層の構造をもつ。層と層の間にはシナプス結合荷重 W があり、ニューロン同士を繋ぐ役割を果たしている。シナプス結合荷重のことを重みと呼ぶ。ニューラルネットワークの中間層は多数重ねることができる。中間層が多層構造になったニューラルネットワークのことを深層学習 (Deep Learning) と呼ぶ。

2.2 RNN

RNN (Recurrent Neural Network) とは、ある時刻における中間層の出力を次の層の入力とすることができる再帰的構造を持つニューラルネットワークモデル

のことである。このニューラルネットワークモデルは、David E. Rumelhart^[4]の研究を元にして開発されたものである。過去の情報を扱うことができるように、前の時刻での中間層の出力を用いることにより、それ以降の出力に影響を与えていく。これにより、過去の情報を保持していくことができるため、時系列データを扱うことが可能になっている。データの形状は、 $\mathbf{x}(1), \dots, \mathbf{x}(T)$ という T 個のデータが入力データ群となっている。RNN では、入力層と中間層の間の重み U 、時刻 t での入力を $x(t)$ 、 f, g を活性化関数として中間層の出力値 $s(i)$ とネットワークの出力 $y(t)$ は以下のように示される。

$$y(t) = f(U\mathbf{x}(t) + W\mathbf{s}(t-1)) \quad (1)$$

入力 x と $s(t-1)$ により $s(t)$ が更新されていく。

$$y(t) = g(W\mathbf{s}(t)) \quad (2)$$

2.3 LSTM

LSTM (Long Short-Term Memory) は、RNN の欠点を補うようにして構築されたニューラルネットワークモデルである。これは、1997年に Sepp Hochreiter や Jürgen Schmidhuber が発表したモデルであり、時系列データを扱うことができる^[5]。これは、長期記憶 (Long Term Memory) と短期記憶 (Short Term Memory) の二つを組み合わせられており、RNN では時刻が離れているデータ間の依存関係を学習することが難しかったという課題を克服するために開発されたモデルである。近い過去を扱う短期記憶と遠い過去を扱う長期記憶が可能であるため、時系列データが長い場合には有効な場合が多い。LSTM は RNN の中間層を LSTM block というメモリと入力ゲート及び忘却ゲート、出力ゲートの三つのゲートを持つブロックに置換することで実現されている。メモリは入

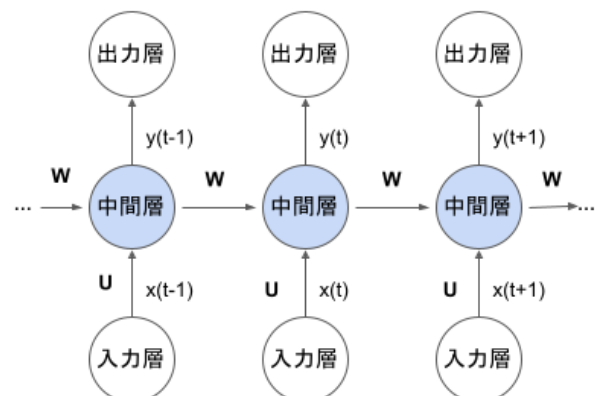


図1 RNN モデル図

力の依存性を記憶している。入力ゲートはメモリへ渡す入力を調整し、忘却ゲートはメモリの値を破棄する量を調整、出力ゲートはメモリの値を活性化関数に渡す量を調整する役割を持っている。

$$\mathbf{f}(t) = \sigma_g(U_f \mathbf{x}(t) + W_f \mathbf{h}(t-1)) \quad (3)$$

$$\mathbf{i}(t) = \sigma_g(U_i \mathbf{x}(t) + W_i \mathbf{h}(t-1)) \quad (4)$$

$$\mathbf{o}(t) = \sigma_g(U_o \mathbf{x}(t) + W_o \mathbf{h}(t-1)) \quad (5)$$

$$\mathbf{c}(t) = \mathbf{f}(t) \odot \mathbf{c}(t-1) + \mathbf{i}(t) \odot \sigma_c(U_c \mathbf{x}(t) + W_c \mathbf{h}(t-1)) \quad (6)$$

$$y(t) = \mathbf{o}(t) \odot \sigma_h(\mathbf{c}(t)) \quad (7)$$

以上より、活性化関数 σ 、入力層と中間層間の重み U 、中間層と出力層間の重み W を用いて忘却ゲート $\mathbf{f}(t)$ や入力ゲート $\mathbf{i}(t)$ 、活性化ベクトル $\mathbf{o}(t)$ 、中間層の出力値 $\mathbf{h}(t)$ 、ネットワークの出力値 $y(t)$ を定義している。 \odot はアダマール積を表しており、ベクトルの要素ごとに積を求める演算子である。

3. 電力使用量予測実験

3.1 対象データ

本研究では、東京電力パワーグリッド社が提供している電力使用量のデータを用い、2016年4月から2022年3月までの期間を用いる。時間ごとのデータは24(時間) × 365(日) × 6(年)で52560個存在している。日ごとのデータは24時間ごと、週ごとのデータは日ごとのデータを7日ごとにまとめて使用

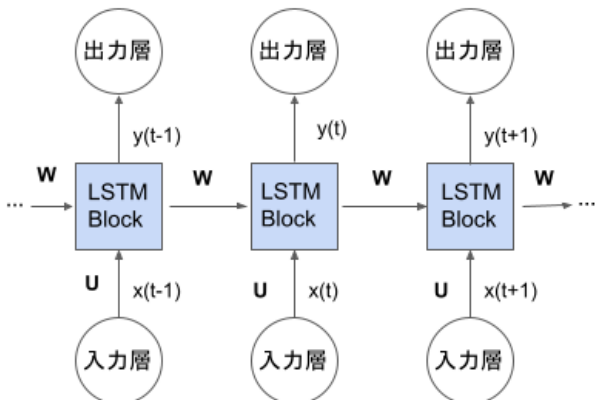


図2 LSTMモデル図

した。電力使用量のデータの一部を図3に示す。2016年11月21日から2週間分の電力使用量の波形を示している。縦軸は電力使用量の 10^4 kW、横軸は時間を表している。図は、日ごとのデータを示しており、一日の間で規則的な波が現れている。図より、日中は電力使用量が増え、夜間は減っていることが読み取れる。本実験ではこのような電力使用量のデータを用いた。

3.2 異常値検出

電力使用量が、トレンドから外れるような急激な変化点を以降では異常値と呼ぶ。異常値は、台風や豪雪などの、異常気象による影響により現れる傾向が高い。そのような異常気象のときの電力使用量を予測することが大きな課題であるが、これまでの研究では精度の高い予測ができていない。用いた電力使用量には、異常値が多く含まれており、そのような点についても高い予測精度を示すことが可能になれば、適切な電力調整や化石燃料調達に役立つ。本実験では、実測データから異常値検出を行うが、実際の予測では、天気や降水量などの電力使用量に関わる他要因も含め、異常値が現れそうな位置を検出することを想定している。

時系列データの平滑化には、移動平均という手法が存在し、大きく分類すると単純移動平均と加重移動平均、指数移動平均の3種類に分類できる。移動平均とは、金融分野の分析や気象の分析に用いられる手法のことである。

単純移動平均とは、SMA (Simple Moving Average) と呼ばれ、直近の n 個のデータを平均をとり、単純な重み付けをすることなく、平滑化を行う手法のことである。しかし、重み付けがないため、過去のデータの影響を受けてしまう。

加重移動平均とは、WMA (Weighted Moving Average) と呼ばれ、重みを一定量ずつ線形に減らすことで平滑化を行う手法のことである。直近のデータの

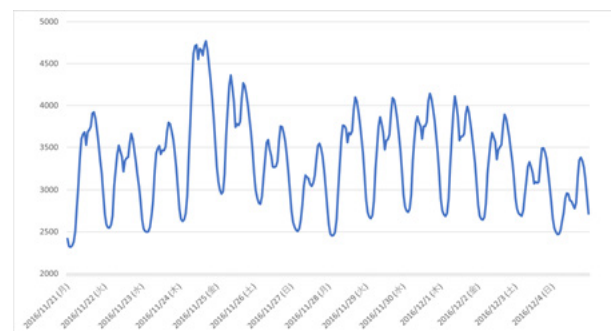


図3 電力使用量の元データ

重みを n として、その一つ前のデータの重みを $n-1$ とする。このように重みを 0 まで小さくしていくことで古いデータの重みを減らし、古いデータの影響を小さくする手法である。

指数移動平均とは、EMA (Exponential Moving Average) と呼ばれ、加重移動平均では線形に減らした重みを指数関数的に減少させる手法のことである。単純移動平均と比べて直近のデータに影響を受けやすいという特徴がある。

まず、指数移動平均による時系列データの平滑化を行った。期間 14 日ごとのデータを用いて移動標準偏差 (Moving Standard Deviation) を計算し、同様に 14 日ごとの指数移動平均を計算する。指数移動平均の計算結果が移動標準偏差から 1 倍以上大きい値を異常値として扱った。図 4 に異常値を検出した結果を示す。青色が元の波形、橙色の波形が EMA、橙色の点が異常値を表している。図から読み取れるように、トレンドから大きく外れているところが異常値となっている。評価は全区間と異常値のみを集めた区間の 2 区間に分けて行う。それぞれの区間について RMSE や分散の指標により評価を行う。

3.3 RNN による予測手法

本研究では、RNN に入力として時間ごとや日ごと、週ごとの電力使用量のデータを与える。このデータから時間ごと、日ごと、週ごとの電力使用量の予測を行った。用いたデータを、Train data と Test data が 8:2 になるように分割し、入力データとした。入力系列の長さは 12 とした。そのため、時間ごとの予測なら 12 時間の電力使用量のデータを入力し、次の 1 時間の電力使用量を予測するような予測方法である。同様に、日ごと、週ごとの予測も行う。最適化アルゴリズムは Adam を用い、学習率の初期値は 0.001、損失関数として平均二乗誤差を用いた。活性化関数は ReLU、バッチサイズは 64 とし、epoch 数は validation loss が 10 回以上改善しなくなるまで繰り返すこととした。層数と中間層の素子数を変化させることで予測を行った。層数は 1 層、2 層、3 層を変化させ、素子数は 100 個、200 個、300 個を変化させることによって予測精度を調べた。RMSE と分散を用いることによって、予測した電力使用量の予測結果の評価を行った。また、RMSE の式は以下で定義される。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (8)$$

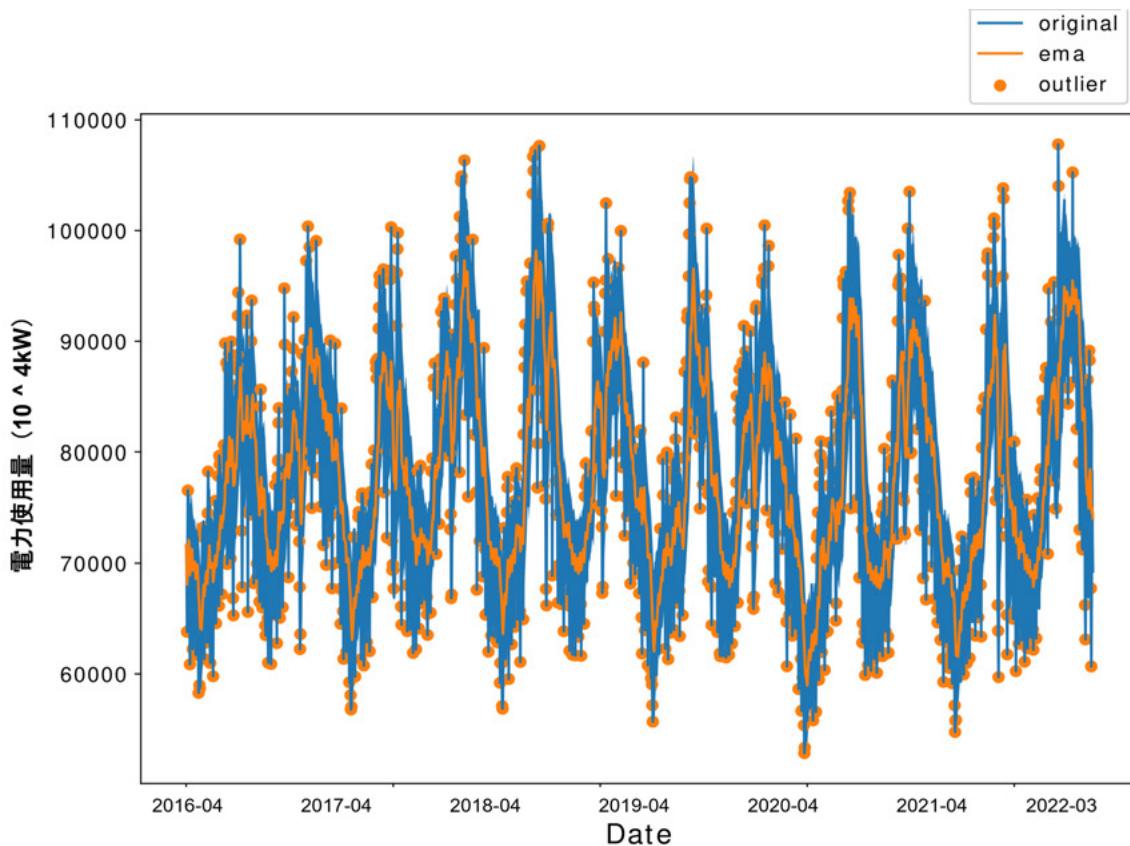


図 4 異常値検出結果

n はデータ数, y と \hat{y} はそれぞれ入力値と予測値である。

3.4 LSTM による予測手法

RNN のときと同様に, LSTM に入力として時間ごとや日ごと, 週ごとの電力使用量のデータを与えた。このデータから時間ごと, 日ごと, 週ごとの電力使用量の予測を行った。用いたデータを, Train data と Test data が 8:2 になるように分割し, 入力データとした。入力系列の長さは 12 とした。そのため, 時間ごとの予測なら 12 時間の電力使用量のデータを入力した。同様に, 日ごと, 週ごとの予測も行う。最適化アルゴリズムは Adam を用い, 学習率の初期値は 0.001, 損失関数として平均二乗誤差を用いた。活性化関数は ReLU, バッチサイズは 64 とし, epoch 数は validation loss が 10 回以上改善しなくなるまで繰り返すこととした。層数と中間層の素子数を変化させることで予測を行った。層数は 1 層, 2 層, 3 層を変化させ, 素子数は 100 個, 200 個, 300 個を変化させることによって予測精度を調べた。RMSE と分散を用いることによって, 予測した電力使用量の予測結果の評価を行った。

表 1 RNN を用いた時間ごとの予測

| RMSE / 分散 | Test (全区間) |
|-----------|-----------------|
| 1 層 (75) | 0.0177 / 0.0323 |
| 1 層 (100) | 0.0168 / 0.0349 |
| 1 層 (125) | 0.0207 / 0.0319 |
| 2 層 (100) | 0.0173 / 0.0338 |
| 2 層 (125) | 0.0172 / 0.0329 |
| 2 層 (150) | 0.0176 / 0.0338 |
| 3 層 (100) | 0.0184 / 0.0320 |
| 3 層 (125) | 0.0163 / 0.0334 |
| 3 層 (150) | 0.0180 / 0.0338 |

3.5 結果

電力使用量の元データ, 学習データとテストデータに対する予測結果を一例として図 5 に示す。電力使用量の元データは緑色, 学習データの予測は青色, テストデータに対する予測は橙色で示している。

RNN と LSTM を用いて, 時間ごと, 日ごと, 週ごとの電力使用量を予測した結果を表 1 に示す。これは, RMSE と分散を用いて, Test data について予測結果を評価した表である。

RNN を用いた時間ごとの予測結果は, 3 層, ニューロン数 125 個のときが最も精度の良い予測となつて

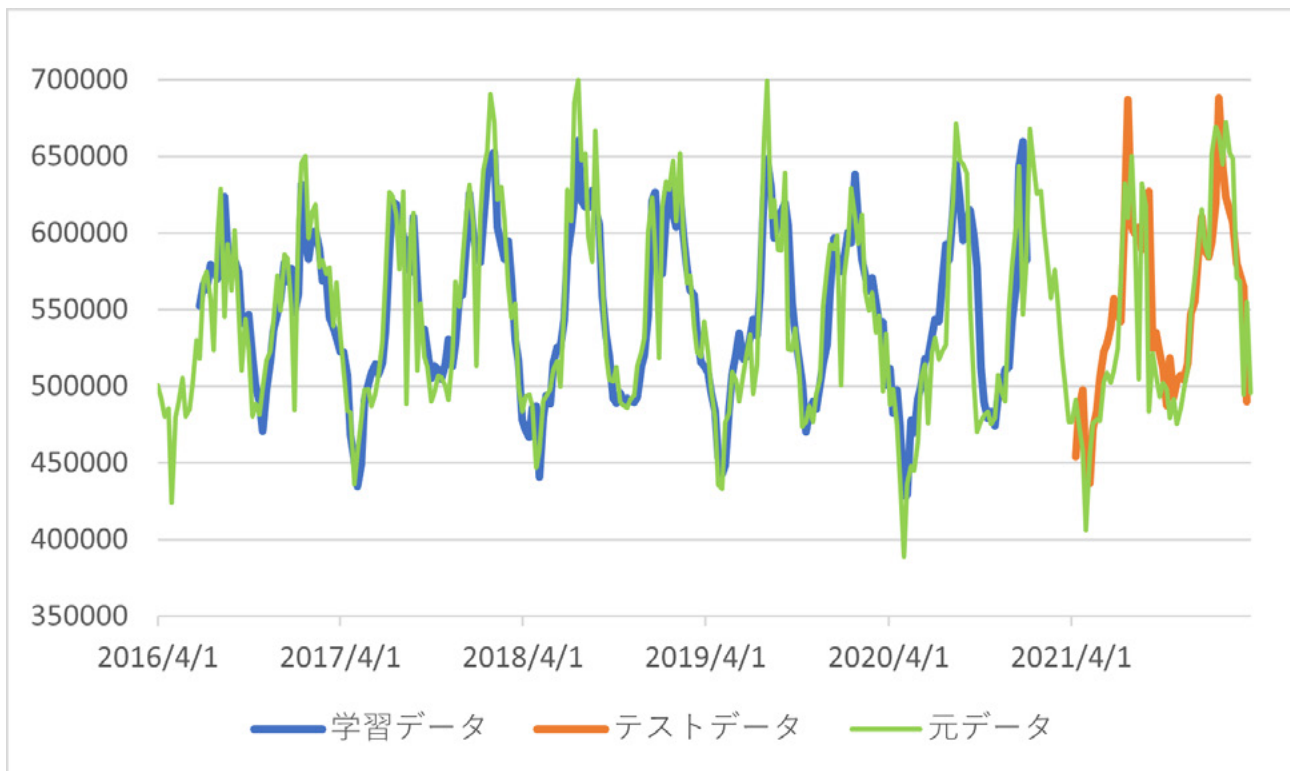


図 5 モデルの予測

表 2 LSTM を用いた時間ごとの予測

| RMSE / 分散 | Test (全区間) |
|-----------|-----------------|
| 1 層 (75) | 0.0158 / 0.0323 |
| 1 層 (100) | 0.0148 / 0.0320 |
| 1 層 (125) | 0.0176 / 0.0325 |
| 2 層 (100) | 0.0172 / 0.0316 |
| 2 層 (125) | 0.0134 / 0.0330 |
| 2 層 (150) | 0.0165 / 0.0323 |
| 2 層 (100) | 0.0163 / 0.0344 |
| 3 層 (125) | 0.0139 / 0.0333 |
| 3 層 (150) | 0.0148 / 0.0349 |

表 3 区間ごとの予測精度 (時間)

| RMSE/分散 | 異常値以外の区間 | 異常値区間 |
|---------|-----------------|-----------------|
| RNN | 0.0186 / 0.0321 | 0.0148 / 0.0377 |
| LSTM | 0.0148 / 0.0309 | 0.0133 / 0.0373 |

いた。時間ごとのデータはデータ数が多く、層を重ねることで複雑な予測に対応しているためこの結果となっている。

表 2 より、LSTM を用いた時間ごとの予測結果は、2 層、ニューロン数 125 個のときが最も精度の良い予測となっていた。RNN のときと同様に、層を重ねて複雑な予測に対応しているためこの結果となっている。

また、時間ごとの異常値区間の予測精度を表 3 に示す。

RMSE の結果は異常値区間の方が低くなっていることが読み取れる。分散の値は大きくなっており、精度が高くないことがわかる。RNN と LSTM では LSTM の方が区間別にみても精度が高くなっている。

表 4 より、RNN を用いた日ごとの予測結果は、1 層、ニューロン数 125 個のときが最も精度の良い予測となっていた。時間ごとのデータと比べて、日ごとのデータはデータ数が少ないため、層数は 1 層が一番良くなっているが、1 層から 2 層までで精度に大きな差は生まれていない結果となった。

表 5 より、LSTM を用いた日ごとの予測結果は、2 層、ニューロン数 125 個のときが最も精度の良い予測となっていた。時間ごとの予測結果と同じ結果となった。また、RNN と LSTM どちらの予測も、時間ごとの結果と比べて RMSE や分散の値が大きくなっていることがわかる。

また、日ごとの異常値区間の予測精度を表 6 に示す。

表 4 RNN を用いた日ごとの予測

| RMSE / 分散 | Test (全区間) |
|-----------|---------------|
| 1 層 (100) | 0.095 / 0.044 |
| 1 層 (125) | 0.091 / 0.041 |
| 1 層 (150) | 0.093 / 0.042 |
| 2 層 (100) | 0.094 / 0.042 |
| 2 層 (125) | 0.093 / 0.041 |
| 2 層 (150) | 0.098 / 0.041 |
| 3 層 (100) | 0.107 / 0.042 |
| 3 層 (125) | 0.098 / 0.042 |
| 3 層 (150) | 0.099 / 0.041 |

表 5 LSTM を用いた日ごとの予測

| RMSE / 分散 | Test (全区間) |
|-----------|---------------|
| 1 層 (100) | 0.103 / 0.042 |
| 1 層 (125) | 0.098 / 0.040 |
| 1 層 (150) | 0.111 / 0.042 |
| 2 層 (100) | 0.128 / 0.042 |
| 2 層 (125) | 0.097 / 0.043 |
| 2 層 (150) | 0.109 / 0.043 |
| 3 層 (100) | 0.112 / 0.042 |
| 3 層 (125) | 0.109 / 0.044 |
| 3 層 (150) | 0.130 / 0.038 |

日ごとのデータの区間ごとの予測精度では、RNN と LSTM ともに異常値区間の RMSE が低くなっていた。分散は、異常値区間が大きくなり、予測精度が高くないことがわかった。全区間で高い精度を示したのは RNN だが、区間ごとに計算しても同じ結果になったと言える。

表 7 より、RNN を用いた週ごとの予測結果は、3 層、ニューロン数 125 個のときが最も精度の良い予測となっていた。データ数が一番少ない週ごとのデータであるが、予測結果は 3 層、ニューロン数 125 個であった。しかし、1 層から 3 層までの予測結果では大きな差はない結果となっていた。

表 8 より、LSTM を用いた週ごとの予測結果は、3 層、ニューロン数 125 個のときが最も精度の良い予測となっていた。1 層、2 層、3 層で結果を比べると 3 層のときの結果が 1 層、2 層に比べて精度が高くなっていた。

また、週ごとの異常値区間の予測精度を表 9 に示す。

週ごとのデータの区間ごとの予測精度は、RNN と LSTM ともに異常値区間の方が RMSE の値が大き

表 6 区間ごとの予測精度 (日)

| RMSE/分散 | 異常値以外の区間 | 異常値区間 |
|---------|---------------|---------------|
| RNN | 0.088 / 0.041 | 0.109 / 0.051 |
| LSTM | 0.088 / 0.043 | 0.116 / 0.055 |

表 7 RNN を用いた週ごとの予測

| RMSE / 分散 | Test (全区間) |
|-----------|---------------|
| 1 層 (100) | 0.194 / 0.061 |
| 1 層 (125) | 0.172 / 0.065 |
| 1 層 (150) | 0.187 / 0.065 |
| 2 層 (100) | 0.178 / 0.059 |
| 2 層 (125) | 0.175 / 0.064 |
| 2 層 (150) | 0.182 / 0.062 |
| 3 層 (100) | 0.180 / 0.062 |
| 3 層 (125) | 0.157 / 0.062 |
| 3 層 (150) | 0.181 / 0.062 |

なり、予測精度が低くなっていた。RNN では分散は、僅かに異常値区間の方が精度が低く、LSTM では僅かに異常値区間の方が精度が高いという結果であった。

時間ごと、日ごと、週ごとの電力の予測を RNN と LSTM を用いて行ってきたが、予測精度が高かったものは全てニューロン数が 125 個であることがわかった。よって、電力使用量の予測にはどちらのモデルを用いるときでもニューロン数を 125 個に設定することで高い精度が得られる。

4. 考察

時間ごとの予測結果は RNN の場合では 3 層、ニューロン数 125 個のとき、LSTM では 2 層、ニューロン数が 125 個のときが最も精度の良い予測となっている。RNN は膨大な期間のデータを記憶することには向いていないため、少ない層数では予測できなかったものだと考えられる。LSTM では、RNN とは異なり、長期記憶が可能になっていることから、層数は 2 層で予測できている。RNN より LSTM の方が全体的に RMSE、分散どちらの数値も小さくなっており、LSTM の方が予測精度が高くなっていた。また、異常値については、分散が大きくなっているため、異常値区間の電力使用量の予測は精度が高くないことがわかった。

日ごとの予測結果は、RNN の場合は 1 層、ニューロン数 125 個のとき、LSTM では 2 層、ニューロン

表 8 LSTM を用いた週ごとの予測

| RMSE / 分散 | Test (全区間) |
|-----------|---------------|
| 1 層 (100) | 0.199 / 0.048 |
| 1 層 (125) | 0.167 / 0.050 |
| 1 層 (150) | 0.234 / 0.051 |
| 2 層 (100) | 0.214 / 0.049 |
| 2 層 (125) | 0.201 / 0.051 |
| 2 層 (150) | 0.223 / 0.045 |
| 3 層 (100) | 0.166 / 0.048 |
| 3 層 (125) | 0.162 / 0.052 |
| 3 層 (150) | 0.190 / 0.048 |

表 9 区間ごとの予測精度 (週)

| RMSE/分散 | 異常値以外の区間 | 異常値区間 |
|---------|---------------|---------------|
| RNN | 0.171 / 0.071 | 0.211 / 0.072 |
| LSTM | 0.179 / 0.059 | 0.223 / 0.058 |

数 125 個のときが最も精度の高い予測となっている。RNN のときは、時間ごとのデータよりもデータ数が少ないために層数が 1 層で高い精度の予測ができたと考えられる。LSTM では 2 層のときが精度が高くなっているが、1 層の結果と値に大きな差はなく、層数は RNN と LSTM どちらの場合も複雑に重ねる必要がない。それぞれの結果を比較すると、RNN の方が全体的に RMSE と分散が小さくなっており、精度が高いといえる。異常値区間については予測精度が低くなっていた。時間ごとのデータの予測と同様に、日ごとのデータの予測も異常値区間の予測精度が高くないことがわかった。

週ごとの予測結果は、RNN の場合は 3 層、ニューロン数 125 個のとき、LSTM では 3 層、ニューロン数 125 個のときが最も精度の高い予測となっている。RNN、LSTM とともに層数が多くなっている。週ごとのデータでは、1 時間の電力使用量のデータを 24 時間分合わせ、さらに 7 日分合わせて作成している。天気や気温などが不規則なことに関連して、1 週間の電力使用量は複雑なデータとなっている可能性が高い。データ数が少ないため層数は少なくとも予測できるとは考えられるが、週ごとのデータが複雑なデータとなっていることにより層数は RNN と LSTM どちらも 3 層が予測精度が高くなったものだと考えられる。異常値区間については、全区間に比べてかなり予測精度が悪くなっていることがわかった。これもデータ数が少ないこと、データが複雑なことにより精度

が下がったものだと考えられる。

データ量によって RNN と LSTM どちらのモデルを用いるか、層数の設定は変える必要があることがわかった。ニューロン数については、125 個がほとんどの結果において高い精度を示していた。しかし、異なるニューロン数の結果と比べて大きく精度が高くなっているとはいえない。今後は、用いた電力使用量のデータ量より長期のデータを用いて実験を行い、これが適切なニューロン数であるか検討する必要がある。

結果より、データ量が非常に多くなるほど LSTM が予測に向いており、反対にデータ量の少ない場合は RNN の方が予測に向いている傾向があることがわかった。用いるデータの量に合わせて適切なモデルとそのモデル構造を選択することで、より精度の高い電力使用量の予測の実現に繋がる。また、異常値区間については、精度が低くなっていたため、電力使用量以外の他要因を用いる必要があると考えられる。そうすることにより、天気や降水量など、電力使用量の変化に影響する他要因からも予測が行えるようになるため、集合型のマルチモーダル入力による学習を行う必要がある。集合型のマルチモーダル学習とは、電力使用量のみから電力使用量の予測を行うモデルと、天気から電力使用量の予測を行うモデルや降水量から電力使用量の予測を行うモデルをそれぞれ作り、それらの予測結果の多数決をとる手法である。ここでいう予測結果は、前の時刻から次の時刻にかけての電力使用量の変化量を予測することを表す。このように予測することで急激に電力使用量が増える時刻を予測できるようになると考えられる。

5. おわりに

本研究では、発電所の稼働状態を調整することや適切な化石燃料の調達のための電力使用量の予測を行った。予測結果は RMSE や分散を用いて精度を評価した。RNN や LSTM を用いた予測では、時間ごと、日ごと、週ごとについてそれぞれ適切なモデルが存在し、電力使用量のトレンドを予測することができるようになった。今後は全区間においてより精度の高い精度を追求していく。また、異常値区間の予測精度については、全区間の予測精度に比べて低くなっていた。異常値は、天気や降水量、気温などの要因から起こる可能性が高い。今回の実験では、電力使用量のみを単入力として予測をしたため、異常値の予測には対応できていなかったと考えられる。

今後の課題は、電力使用量の単入力での予測精度の向上を図ることである。また、電力使用量のみを入

力データとして扱うのではなく、天気や降水量などの他要因もマルチモーダルな入力として扱う深層学習モデルを構築することである。これにより、電力使用量のトレンドのみではなく、他要因により現れる異常値についても詳細な予測ができるようになる。

参考文献

- [1] 市村匠. 鎌田真. リカレント構造適応型 *Deep Belief Network* による時系列データの学習. 計測自動制御学会論文集, Vol.54, No.8, pp.628-639, 2018.
- [2] 村上朋子. 松尾雄司, 永富悠. 有価証券報告書を用いた火力・原子力発電コスト構造の分析. *Journal of Japan Society of Energy and Resources*, Vol.33, No.5, pp.1-2, 2012.
- [3] 浅川伸一. *python* で体験する深層学習. pp.302-303, コロナ社, 2016.
- [4] G.E.Hinton D.E.Rumelhart and R.J.Williams. *Learning representations by back-propagating errors*. *Nature*, Vol.323, pp.533-536, 1986.
- [5] J.Schmidhuber. S.Hochreiter. *Long short-term memory*. *Neural Computation*, Vol.9, No.8, pp.1735-1780, 1997.